MSCA 31006 Assignment2

Duo Zhou

10/18/2020

Instructions:

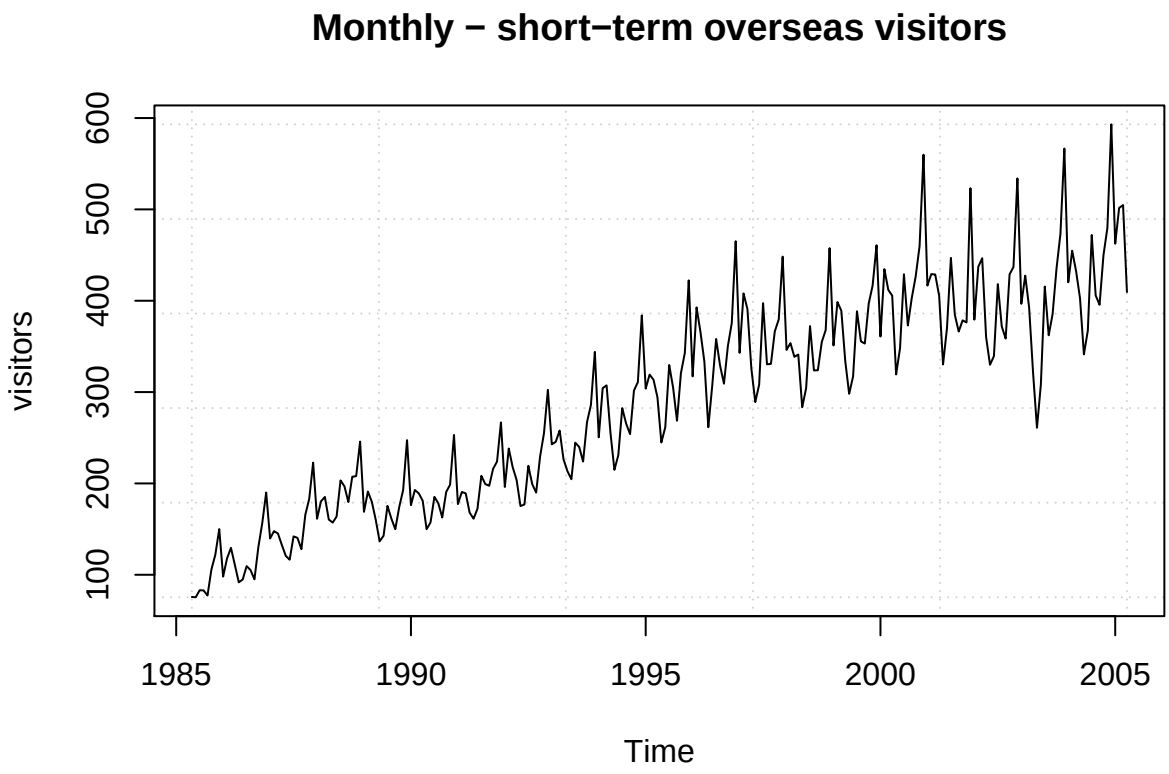
- Total number of points is 30. The assignmnet's final grade will be multiplied by 1/6 to calculate its weight on the final grade.
- Mark the question number and your final answer clearly (use a textbox.)
- Remember to show and explain your work (If you can't explain it, you don't understand it.)
- Please submit your solution through Canvas.

For this exercise, use the monthly Australian short-term overseas visitors data, May 1985–April 2005 which is also available in the 'expsmooth' library (Data set: visitors.)

(5 points) Question 1:

Load the visitors.rda dataset, make a time plot of your data and describe the main features of the series.

```
data(visitors)
plot(visitors, main="Monthly - short-term overseas visitors",panel.first = grid())
```



Based on the TS plot above, we can see that there is an increasing trend(means of TS increases over time), which implies that the time series is no stationary. There is a strong seasonality in the time series, but no cyclic pattern.

(5 points) Question 2: What is the appropriate Holt-Winters method for this data set (multiplicative / additive)? why?

Since the amplitude of the seasonal pattern is proportional to the average level within the season(changes over time), we can say that the appropriate Holt-Winters seasonal model for the data set is multiplicative method.

(10 points) Question 3 Use the hw() function to forecast the next 15 months using Holt-Winters' methods. Experiment

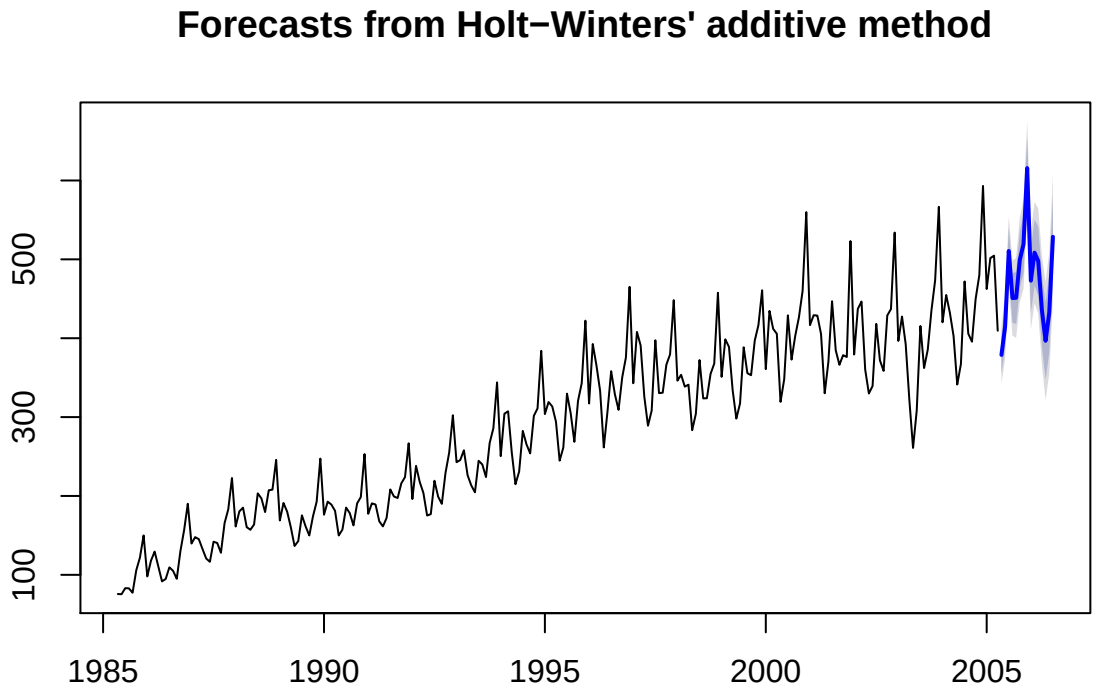
with the following methods

```
lt_ad <- hw(visitors, h=15, seasonal="additive", damped=FALSE,
            exponential = FALSE)
summary(lt_ad)
```

Linear trend with additive seasonality

```
##
## Forecast method: Holt-Winters' additive method
##
## Model Information:
## Holt-Winters' additive method
##
## Call:
##  hw(y = visitors, h = 15, seasonal = "additive", damped = FALSE,
##
##  Call:
##      exponential = FALSE)
##
##   Smoothing parameters:
##     alpha = 0.4819
##     beta  = 1e-04
##     gamma = 0.3245
##
##   Initial states:
##     l = 104.4488
##     b = 1.4956
##     s = -16.1688 14.864 24.4611 -6.1019 90.3471 23.995
##            8.7615 -33.1694 -17.4486 10.1896 -43.028 -56.7016
##
##   sigma:  18.65
##
##      AIC     AICc      BIC
## 2737.200 2739.957 2796.371
##
## Error measures:
##                      ME     RMSE     MAE        MPE     MAPE      MASE      ACF1
## Training set 0.0515815 18.01758 13.7496 -0.1392964 5.413221 0.5077597 0.1379352
##
## Forecasts:
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## May 2005       378.7927 354.8918 402.6936 342.2395 415.3460
## Jun 2005       414.4400 387.9072 440.9729 373.8615 455.0185
## Jul 2005       510.6476 481.7202 539.5751 466.4069 554.8883
## Aug 2005       450.9204 419.7811 482.0598 403.2969 498.5440
## Sep 2005       451.2590 418.0537 484.4642 400.4759 502.0420
## Oct 2005       499.4119 464.2611 534.5627 445.6533 553.1704
## Nov 2005       519.2091 482.2141 556.2042 462.6301 575.7882
## Dec 2005       615.7079 576.9556 654.4603 556.4413 674.9746
## Jan 2006       472.9433 432.5091 513.3775 411.1045 534.7821
## Feb 2006       508.4579 466.4084 550.5075 444.1487 572.7672
## Mar 2006       497.9718 454.3660 541.5776 431.2825 564.6611
## Apr 2006       435.5321 390.4230 480.6412 366.5437 504.5205
## May 2006       396.7569 347.6900 445.8237 321.7156 471.7981
## Jun 2006       432.4042 381.9954 482.8130 355.3106 509.4977
## Jul 2006       528.6118 476.8953 580.3283 449.5182 607.7053
```

```
plot(lt_ad)
```



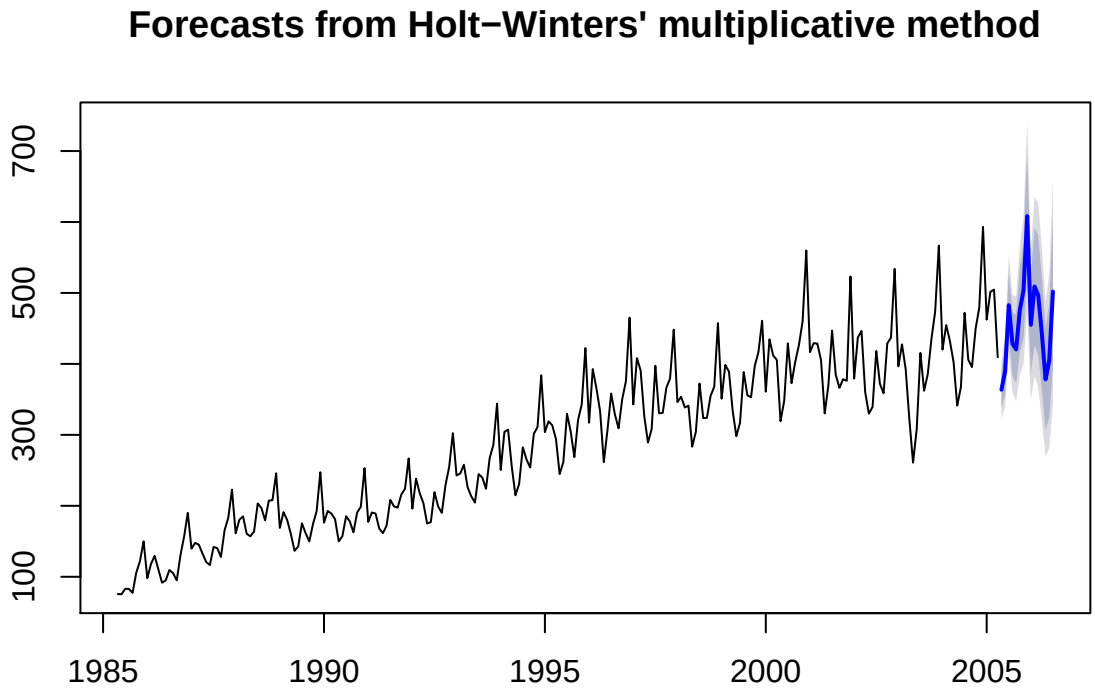

```
lt_mu <- hw(visitors, h=15, seasonal="multiplicative", damped=FALSE,
            exponential = FALSE)
summary(lt_mu)
```

Linear trend with multiplicative seasonality

```
##
## Forecast method: Holt-Winters' multiplicative method
##
## Model Information:
## Holt-Winters' multiplicative method
##
## Call:
##  hw(y = visitors, h = 15, seasonal = "multiplicative", damped = FALSE,
##
##  Call:
##      exponential = FALSE)
##
##   Smoothing parameters:
##     alpha = 0.5653
##     beta  = 0.0215
##     gamma = 5e-04
##
```

```
##   Initial states:
##     l = 91.7613
##     b = 2.4333
##     s = 0.935 1.0545 1.0841 0.9724 1.3037 1.0824
##           1.0258 0.9102 0.9304 1.0521 0.8518 0.7976
##
##   sigma:  0.0565
##
##      AIC     AICc      BIC
## 2628.219 2630.976 2687.390
##
## Error measures:
##                       ME    RMSE      MAE        MPE     MAPE      MASE
## Training set -0.09495709 14.6622 10.97229 -0.3070136 4.188878 0.4051965
##                    ACF1
## Training set 0.07998858
##
## Forecasts:
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## May 2005       363.6434 337.2901 389.9967 323.3395 403.9474
## Jun 2005       389.5974 356.8764 422.3184 339.5550 439.6399
## Jul 2005       482.7709 437.0247 528.5172 412.8081 552.7338
## Aug 2005       428.5001 383.4951 473.5050 359.6709 497.3293
## Sep 2005       420.4548 372.1236 468.7860 346.5386 494.3710
## Oct 2005       475.5963 416.3290 534.8636 384.9547 566.2379
## Nov 2005       503.4598 435.9471 570.9725 400.2080 606.7116
## Dec 2005       608.3779 521.1122 695.6436 474.9165 741.8393
## Jan 2006       455.1525 385.6597 524.6452 348.8725 561.4324
## Feb 2006       509.0590 426.6702 591.4478 383.0562 635.0618
## Mar 2006       496.8101 411.8771 581.7431 366.9162 626.7039
## Apr 2006       441.9586 362.3923 521.5249 320.2724 563.6448
## May 2006       378.2099 306.6932 449.7266 268.8345 487.5852
## Jun 2006       405.1516 324.8788 485.4244 282.3849 527.9183
## Jul 2006       501.9810 397.9885 605.9736 342.9382 661.0239
```

```
plot(lt_mu)
```



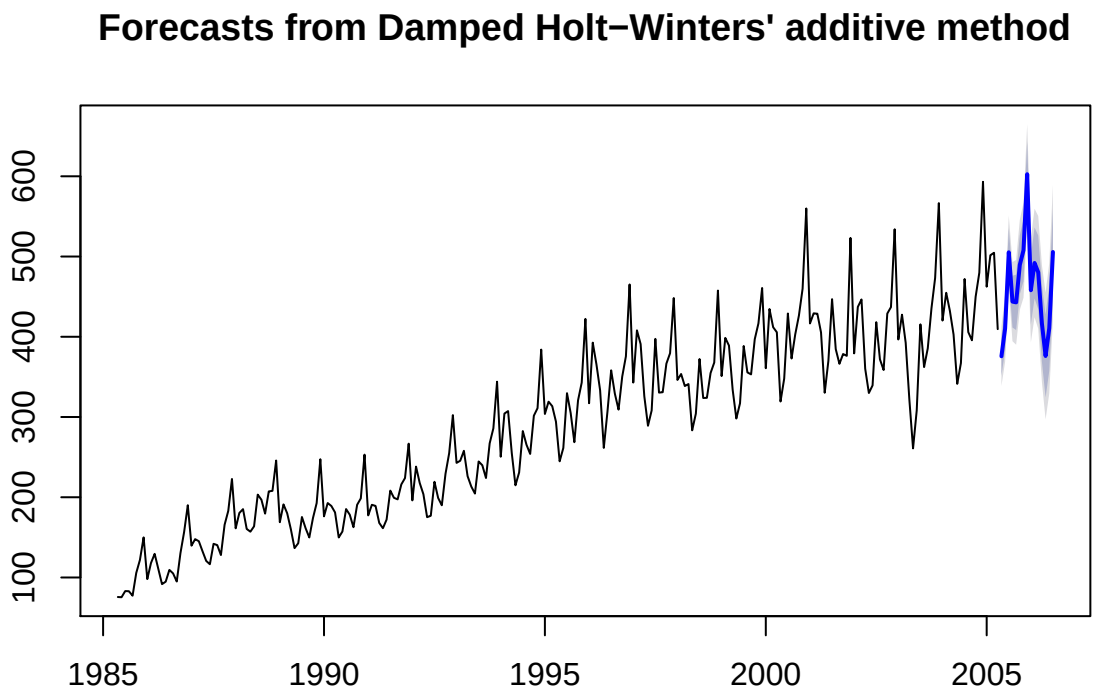

```
lt_ad_dam <- hw(visitors, h=15, seasonal="additive", damped=TRUE,
                exponential = FALSE)
summary(lt_ad_dam)
```

Linear trend with additive seasonality and damping

```
##
## Forecast method: Damped Holt-Winters' additive method
##
## Model Information:
## Damped Holt-Winters' additive method
##
## Call:
##  hw(y = visitors, h = 15, seasonal = "additive", damped = TRUE,
##
##  Call:
##      exponential = FALSE)
##
##   Smoothing parameters:
##     alpha = 0.5108
##     beta  = 1e-04
##     gamma = 0.3251
##     phi   = 0.98
```

```
##
##   Initial states:
##     l = 99.839
##     b = 2.7082
##     s = -17.3293 14.2436 24.2586 -7.1666 90.4445 23.3113
##           7.8487 -28.2181 -18.8684 12.3915 -44.7869 -56.1288
##
##   sigma:  18.8433
##
##      AIC     AICc      BIC
## 2743.077 2746.172 2805.729
##
## Error measures:
##                     ME     RMSE      MAE       MPE     MAPE      MASE      ACF1
## Training set 1.824795 18.16369 13.83982 0.3848574 5.415549 0.5110917 0.1106985
##
## Forecasts:
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## May 2005       375.6699 351.5212 399.8185 338.7376 412.6021
## Jun 2005       410.2359 383.1186 437.3531 368.7636 451.7082
## Jul 2005       505.2769 475.4844 535.0694 459.7132 550.8406
## Aug 2005       443.8172 411.5698 476.0646 394.4990 493.1354
## Sep 2005       442.8545 408.3255 477.3836 390.0469 495.6622
## Oct 2005       489.7055 453.0358 526.3753 433.6240 545.7871
## Nov 2005       507.8249 469.1320 546.5177 448.6492 567.0005
## Dec 2005       602.5978 561.9818 643.2138 540.4810 664.7146
## Jan 2006       457.9759 415.5232 500.4286 393.0501 522.9017
## Feb 2006       491.9928 447.7791 536.2065 424.3738 559.6119
## Mar 2006       480.2909 434.3831 526.1987 410.0810 550.5008
## Apr 2006       416.5024 368.9604 464.0445 343.7932 489.2117
## May 2006       376.0388 324.3791 427.6984 297.0322 455.0454
## Jun 2006       410.5974 357.4793 463.7155 329.3603 491.8345
## Jul 2006       505.6312 451.0932 560.1692 422.2226 589.0399
```

```
plot(lt_ad_dam)
```



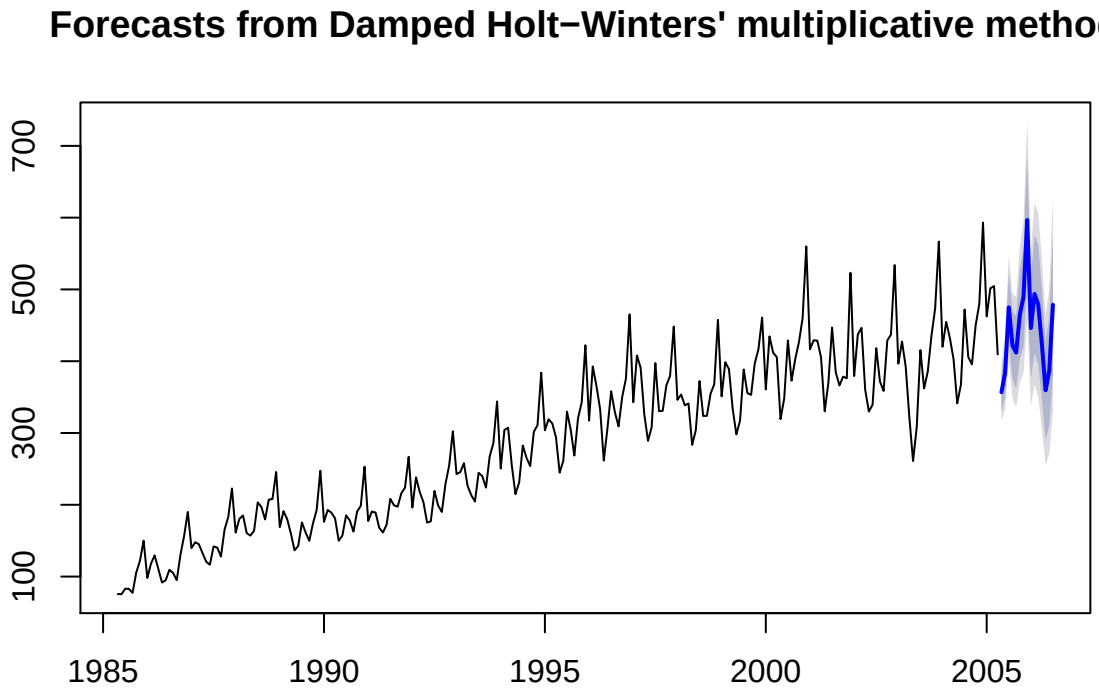

```
lt_mu_dam <- hw(visitors, h=15, seasonal="multiplicative",
                damped=TRUE, exponential = FALSE)
summary(lt_mu_dam)
```

Linear trend with multiplicative seasonality and damping

```
##
## Forecast method: Damped Holt-Winters' multiplicative method
##
## Model Information:
## Damped Holt-Winters' multiplicative method
##
## Call:
##  hw(y = visitors, h = 15, seasonal = "multiplicative", damped = TRUE,
##
##  Call:
##      exponential = FALSE)
##
##   Smoothing parameters:
##     alpha = 0.6668
##     beta  = 0.0043
##     gamma = 1e-04
##     phi   = 0.98
```

```
##
##   Initial states:
##     l = 91.5731
##     b = 2.1794
##     s = 0.9303 1.0531 1.086 0.9822 1.3144 1.0796
##           1.025 0.9094 0.9322 1.05 0.8485 0.7895
##
##   sigma:  0.0568
##
##      AIC     AICc      BIC
## 2628.489 2631.584 2691.140
##
## Error measures:
##                     ME     RMSE      MAE       MPE     MAPE      MASE
## Training set 1.286455 14.41189 10.67154 0.2674105 4.065573 0.3940899
##                      ACF1
## Training set -0.02073956
##
## Forecasts:
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## May 2005       356.7701 330.8024 382.7378 317.0559 396.4842
## Jun 2005       383.6952 350.0518 417.3386 332.2420 435.1484
## Jul 2005       475.1096 427.3402 522.8790 402.0526 548.1666
## Aug 2005       422.1456 374.8352 469.4560 349.7906 494.5006
## Sep 2005       412.0505 361.5166 462.5844 334.7656 489.3354
## Oct 2005       464.8117 403.2339 526.3896 370.6366 558.9869
## Nov 2005       489.8499 420.4189 559.2808 383.6644 596.0354
## Dec 2005       596.7632 506.9344 686.5921 459.3819 734.1446
## Jan 2006       446.1817 375.2749 517.0884 337.7391 554.6242
## Feb 2006       493.6657 411.2359 576.0956 367.6002 619.7313
## Mar 2006       479.0113 395.3086 562.7139 350.9991 607.0234
## Apr 2006       423.3885 346.2252 500.5518 305.3774 541.3996
## May 2006       359.5283 291.3831 427.6736 255.3093 463.7474
## Jun 2006       386.6002 310.5842 462.6163 270.3437 502.8567
## Jul 2006       478.6323 381.2146 576.0499 329.6448 627.6197
```

```
plot(lt_mu_dam)
```



(5 points) Question 4: Use the accuracy() function to compare the Root-Mean-Square-Error (RMSE) values of the forecasts from the various methods. Which do you prefer and why?

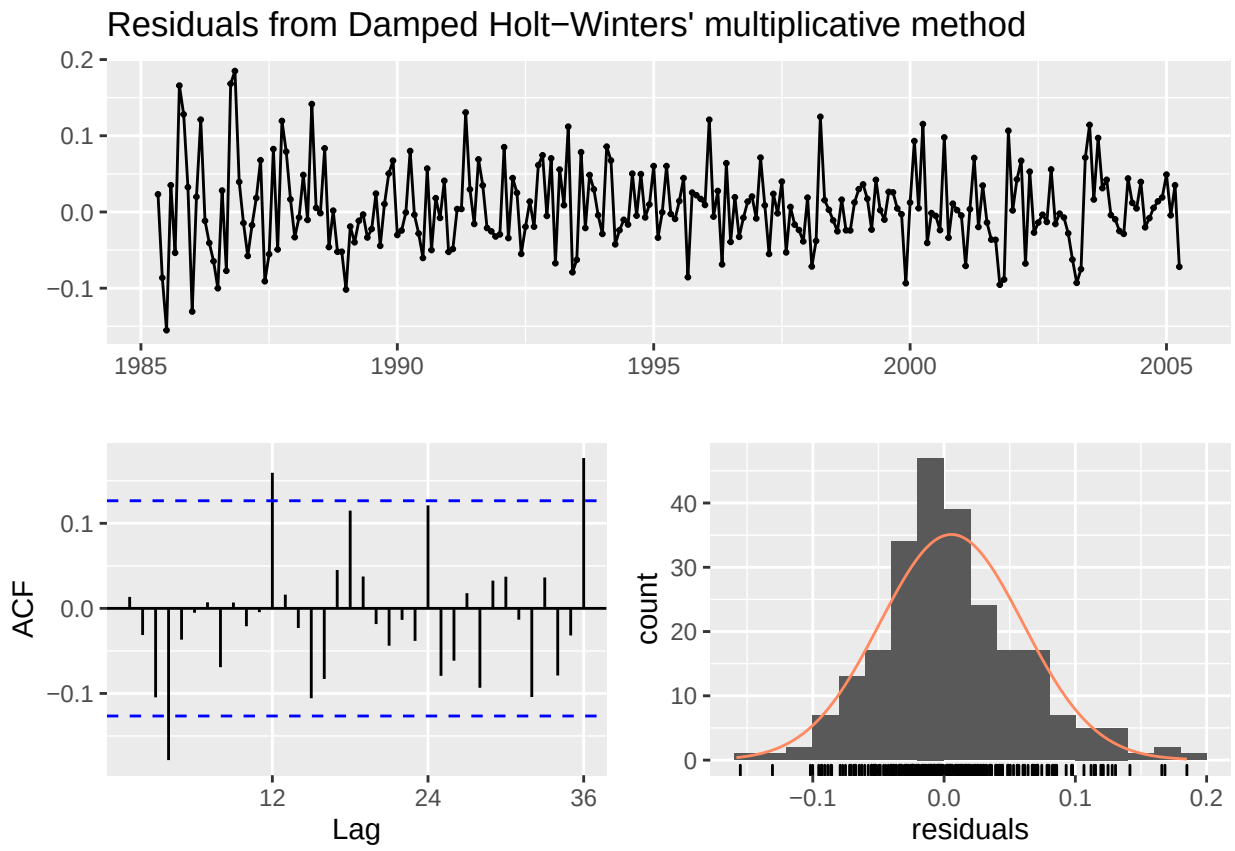
```
cat('The RMSE of Linear trend with additive seasonality model is ',
      accuracy(lt_ad)[2], '\n',
'The RMSE of Linear trend with multiplicative seasonality model is ',
      accuracy(lt_mu)[2], '\n',
'The RMSE of Linear trend with additive seasonality and damping model is ',
      accuracy(lt_ad_dam)[2], '\n',
'The RMSE of Linear trend with multiplicative seasonality and damping model is ',
      accuracy(lt_mu_dam)[2], '\n')
```

```
## The RMSE of Linear trend with additive seasonality model is  18.01758
##  The RMSE of Linear trend with multiplicative seasonality model is  14.6622
##  The RMSE of Linear trend with additive seasonality and damping model is  18.16369
##  The RMSE of Linear trend with multiplicative seasonality and damping model is  14.41189
```

Linear trend with multiplicative seasonality and damping model has the smallest RMSE. It is the most prefered model.

(5 points) Question 5: Use the checkresiduals() function to check that the residuals from the best model look like white noise and provide a summary of the model's smoothing parameters using the summary() function.

```
checkresiduals(lt_mu_dam)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from Damped Holt-Winters' multiplicative method
## Q* = 33.146, df = 7, p-value = 2.487e-05
##
## Model df: 17.   Total lags used: 24
```

Based on the ACF plot, we can see that there are still autocorrelations in lag 3 and lag 12 in the residuals. Box-Ljung test also supports that H0, there is no serial autocorrelation in the residuals at any lag k, should be rejected. The residuals from the best model we chose does not look like a white noise.

We can say that Linear trend with multiplicative seasonality and damping method still left some serial correlations(information not caused by randomness) out of the model.

```
summary(lt_mu_dam)
```

```
##
## Forecast method: Damped Holt-Winters' multiplicative method
##
## Model Information:
## Damped Holt-Winters' multiplicative method
```

```
##
## Call:
##  hw(y = visitors, h = 15, seasonal = "multiplicative", damped = TRUE,
##
##  Call:
##      exponential = FALSE)
##
##   Smoothing parameters:
##     alpha = 0.6668
##     beta  = 0.0043
##     gamma = 1e-04
##     phi   = 0.98
##
##   Initial states:
##     l = 91.5731
##     b = 2.1794
##     s = 0.9303 1.0531 1.086 0.9822 1.3144 1.0796
##            1.025 0.9094 0.9322 1.05 0.8485 0.7895
##
##   sigma:  0.0568
##
##      AIC     AICc      BIC
## 2628.489 2631.584 2691.140
##
## Error measures:
##                     ME     RMSE      MAE       MPE     MAPE      MASE
## Training set 1.286455 14.41189 10.67154 0.2674105 4.065573 0.3940899
##                      ACF1
## Training set -0.02073956
##
## Forecasts:
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## May 2005       356.7701 330.8024 382.7378 317.0559 396.4842
## Jun 2005       383.6952 350.0518 417.3386 332.2420 435.1484
## Jul 2005       475.1096 427.3402 522.8790 402.0526 548.1666
## Aug 2005       422.1456 374.8352 469.4560 349.7906 494.5006
## Sep 2005       412.0505 361.5166 462.5844 334.7656 489.3354
## Oct 2005       464.8117 403.2339 526.3896 370.6366 558.9869
## Nov 2005       489.8499 420.4189 559.2808 383.6644 596.0354
## Dec 2005       596.7632 506.9344 686.5921 459.3819 734.1446
## Jan 2006       446.1817 375.2749 517.0884 337.7391 554.6242
## Feb 2006       493.6657 411.2359 576.0956 367.6002 619.7313
## Mar 2006       479.0113 395.3086 562.7139 350.9991 607.0234
## Apr 2006       423.3885 346.2252 500.5518 305.3774 541.3996
## May 2006       359.5283 291.3831 427.6736 255.3093 463.7474
## Jun 2006       386.6002 310.5842 462.6163 270.3437 502.8567
## Jul 2006       478.6323 381.2146 576.0499 329.6448 627.6197
```

Based on the summary of the model we chose, alpha indicates that how much weights were given to the level from the previous data values. Since alpha is 0.6668, we can say that the influence of previous data levels on current level is moderate. Beta is 0.0043, which indicates that there is a trend in the model, but very small. Gamma is 1e-04, which indicates that the change in seasonality is very insignificant. Finally, Phi is 0.98 which is very close to 1. A phi value close to 1 indicates that the trend is very linear(Very little damping is applied to the original linear trend).