# ISYE6501x - WEEK 3 HW

#### Vivian Peng

Question 7.1 Describe a situation or problem from your job, everyday life, current events, etc., for which exponential smoothing would be appropriate. What data would you need? Would you expect the value of α (the first smoothing parameter) to be closer to 0 or 1, and why?

Walmart can use exponential smoothing to help forecast their future sales of a variety of products for a store. The data needed would be the daily sales of the product in the store in the past years and the past months. I would expect the value of α to be closer to 1 since the recent demand for a product is likely to have more impact in forecasting the future sales than the randomness in the system.

Question 7.2 Using the 20 years of daily high temperature data for Atlanta (July through October) from Question 6.2 (file temps.txt), build and use an exponential smoothing model to help make a judgment of whether the unofficial end of summer has gotten later over the 20 years. (Part of the point of this assignment is for you to think about how you might use exponential smoothing to answer this question. Feel free to combine it with other models if you’d like to. There’s certainly more than one reasonable approach.)

Note: in R, you can use either HoltWinters (simpler to use) or the smooth package’s es function (harder to use,but more general).

If you use “es”,the Holt-Winters model uses model=”AAM” in the function call (the first and second constants are used “A”dditively, and the third (seasonality) is used “M”ultiplicatively; the documentation doesn’t make that clear).

Hide

Data\_temps= read.csv("temps.txt",sep = "",check.names=FALSE)  
str(Data\_temps)

'data.frame': 123 obs. of 21 variables:  
 $ DAY : Factor w/ 123 levels "1-Aug","1-Jul",..: 2 46 90 101 105 109 113 117 121 6 ...  
 $ 1996: int 98 97 97 90 89 93 93 91 93 93 ...  
 $ 1997: int 86 90 93 91 84 84 75 87 84 87 ...  
 $ 1998: int 91 88 91 91 91 89 93 95 95 91 ...  
 $ 1999: int 84 82 87 88 90 91 82 86 87 87 ...  
 $ 2000: int 89 91 93 95 96 96 96 91 96 99 ...  
 $ 2001: int 84 87 87 84 86 87 87 89 91 87 ...  
 $ 2002: int 90 90 87 89 93 93 89 89 90 91 ...  
 $ 2003: int 73 81 87 86 80 84 87 90 89 84 ...  
 $ 2004: int 82 81 86 88 90 90 89 87 88 89 ...  
 $ 2005: int 91 89 86 86 89 82 76 88 89 78 ...  
 $ 2006: int 93 93 93 91 90 81 80 82 84 84 ...  
 $ 2007: int 95 85 82 86 88 87 82 82 89 86 ...  
 $ 2008: int 85 87 91 90 88 82 88 90 89 87 ...  
 $ 2009: int 95 90 89 91 80 87 86 82 84 84 ...  
 $ 2010: int 87 84 83 85 88 89 94 97 96 90 ...  
 $ 2011: int 92 94 95 92 90 90 94 94 91 92 ...  
 $ 2012: int 105 93 99 98 100 98 93 95 97 95 ...  
 $ 2013: int 82 85 76 77 83 83 79 88 88 87 ...  
 $ 2014: int 90 93 87 84 86 87 89 90 90 87 ...  
 $ 2015: int 85 87 79 85 84 84 90 90 91 93 ...

Hide

print(summary(Data\_temps))

DAY 1996 1997 1998 1999   
 1-Aug : 1 Min. :60.00 Min. :55.00 Min. :63.00 Min. :57.00   
 1-Jul : 1 1st Qu.:79.00 1st Qu.:78.50 1st Qu.:79.50 1st Qu.:75.00   
 1-Oct : 1 Median :84.00 Median :84.00 Median :86.00 Median :86.00   
 1-Sep : 1 Mean :83.72 Mean :81.67 Mean :84.26 Mean :83.36   
 10-Aug : 1 3rd Qu.:90.00 3rd Qu.:88.50 3rd Qu.:89.00 3rd Qu.:91.00   
 10-Jul : 1 Max. :99.00 Max. :95.00 Max. :95.00 Max. :99.00   
 (Other):117   
 2000 2001 2002 2003   
 Min. : 55.00 Min. :51.00 Min. :57.00 Min. :57.00   
 1st Qu.: 77.00 1st Qu.:78.00 1st Qu.:78.00 1st Qu.:78.00   
 Median : 86.00 Median :84.00 Median :87.00 Median :84.00   
 Mean : 84.03 Mean :81.55 Mean :83.59 Mean :81.48   
 3rd Qu.: 91.00 3rd Qu.:87.00 3rd Qu.:91.00 3rd Qu.:87.00   
 Max. :101.00 Max. :93.00 Max. :97.00 Max. :91.00   
   
 2004 2005 2006 2007 2008   
 Min. :62.00 Min. :54.00 Min. :53.00 Min. : 59.0 Min. :50.00   
 1st Qu.:78.00 1st Qu.:81.50 1st Qu.:79.00 1st Qu.: 81.0 1st Qu.:79.50   
 Median :82.00 Median :85.00 Median :85.00 Median : 86.0 Median :85.00   
 Mean :81.76 Mean :83.36 Mean :83.05 Mean : 85.4 Mean :82.51   
 3rd Qu.:87.00 3rd Qu.:88.00 3rd Qu.:91.00 3rd Qu.: 89.5 3rd Qu.:88.50   
 Max. :95.00 Max. :94.00 Max. :98.00 Max. :104.0 Max. :95.00   
   
 2009 2010 2011 2012   
 Min. :51.00 Min. :67.00 Min. :59.00 Min. : 56.00   
 1st Qu.:75.00 1st Qu.:82.00 1st Qu.:79.00 1st Qu.: 79.50   
 Median :83.00 Median :90.00 Median :89.00 Median : 85.00   
 Mean :80.99 Mean :87.21 Mean :85.28 Mean : 84.65   
 3rd Qu.:88.00 3rd Qu.:93.00 3rd Qu.:94.00 3rd Qu.: 90.50   
 Max. :95.00 Max. :97.00 Max. :99.00 Max. :105.00   
   
 2013 2014 2015   
 Min. :56.00 Min. :63.00 Min. :56.0   
 1st Qu.:77.00 1st Qu.:81.50 1st Qu.:77.0   
 Median :84.00 Median :86.00 Median :85.0   
 Mean :81.67 Mean :83.94 Mean :83.3   
 3rd Qu.:88.00 3rd Qu.:89.00 3rd Qu.:90.0   
 Max. :92.00 Max. :95.00 Max. :97.0

Hide

Data\_temps

|  |
| --- |
|  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **DAY**  <fctr> | **1996**  <int> | **1997**  <int> | **1998**  <int> | **1999**  <int> | **2000**  <int> | **2001**  <int> | **2002**  <int> | **2003**  <int> | **2004**  <int> |  |
| 1-Jul | 98 | 86 | 91 | 84 | 89 | 84 | 90 | 73 | 82 |  |
| 2-Jul | 97 | 90 | 88 | 82 | 91 | 87 | 90 | 81 | 81 |  |
| 3-Jul | 97 | 93 | 91 | 87 | 93 | 87 | 87 | 87 | 86 |  |
| 4-Jul | 90 | 91 | 91 | 88 | 95 | 84 | 89 | 86 | 88 |  |
| 5-Jul | 89 | 84 | 91 | 90 | 96 | 86 | 93 | 80 | 90 |  |
| 6-Jul | 93 | 84 | 89 | 91 | 96 | 87 | 93 | 84 | 90 |  |
| 7-Jul | 93 | 75 | 93 | 82 | 96 | 87 | 89 | 87 | 89 |  |
| 8-Jul | 91 | 87 | 95 | 86 | 91 | 89 | 89 | 90 | 87 |  |
| 9-Jul | 93 | 84 | 95 | 87 | 96 | 91 | 90 | 89 | 88 |  |
| 10-Jul | 93 | 87 | 91 | 87 | 99 | 87 | 91 | 84 | 89 |  |

Next

**1**23456

...

13

Previous

1-10 of 123 rows | 1-10 of 21 columns

Hide

temps\_mat <- as.vector(unlist(Data\_temps[,2:21]))  
str(temps\_mat)

int [1:2460] 98 97 97 90 89 93 93 91 93 93 ...

Hide

temps\_mat

[1] 98 97 97 90 89 93 93 91 93 93 90 91 93 93 82 91 96 95  
 [19] 96 99 91 95 91 93 84 84 82 79 90 91 87 86 90 84 91 93  
 [37] 88 91 84 90 89 88 86 84 86 89 90 91 91 90 89 90 91 91  
 [55] 91 84 88 84 86 88 84 82 80 73 87 84 87 89 89 89 91 84  
 [73] 86 88 78 79 86 82 82 78 79 79 78 81 84 84 87 84 79 75  
 [91] 72 64 66 72 84 70 66 64 60 78 70 72 69 69 73 79 81 80  
 [109] 82 66 63 68 79 81 69 73 73 75 75 81 82 82 81 86 90 93  
 [127] 91 84 84 75 87 84 87 84 88 86 90 91 91 89 89 89 90 89  
 [145] 84 87 88 89 89 91 91 89 88 72 80 84 88 89 88 84 84 80  
 [163] 73 80 86 88 88 87 88 91 91 89 89 88 82 79 81 82 84 87  
 [181] 90 90 91 91 88 88 91 93 81 81 82 86 88 84 80 82 86 87  
 [199] 87 88 88 90 88 91 95 89 70 80 82 66 70 64 68 77 86 75  
 [217] 73 75 78 81 82 82 82 80 82 82 79 80 68 63 57 66 64 69  
 [235] 70 70 62 63 62 75 71 57 55 64 66 60 91 88 91 91 91 89  
 [253] 93 95 95 91 91 86 88 87 91 87 90 91 95 91 91 89 91 91  
 [271] 86 88 80 88 89 90 86 86 82 84 86 90 89 89 86 82 87 88  
 [289] 84 86 80 82 86 84 87 90 79 84 87 87 88 90 91 89 90 93  
 [307] 93 91 87 84 77 90 91 89 90 89 79 78 81 84 89 87 87 88  
 [325] 87 82 80 82 82 88 84 81 82 84 87 80 75 75 86 78 77 82  
 [343] 82 73 82 69 72 73 78 78 78 75 79 78 77 78 82 75 73 63  
 [361] 63 72 75 79 79 79 78 82 79 84 82 87 88 90 91 82 86 87  
 [379] 87 82 77 73 81 81 86 82 87 88 90 90 91 93 93 91 93 93  
 [397] 93 93 97 99 96 93 88 89 91 93 93 93 91 90 96 98 97 98  
 [415] 93 93 96 98 98 89 91 91 90 80 82 89 88 90 91 91 84 88  
 [433] 91 84 93 96 96 91 91 77 87 87 87 86 87 89 81 81 82 79  
 [451] 68 79 72 75 78 81 82 78 80 77 71 73 75 84 71 73 71 73  
 [469] 73 72 72 73 70 64 75 73 77 80 71 66 60 64 73 57 59 64  
 [487] 69 75 73 72 75 75 89 91 93 95 96 96 96 91 96 99 96 93  
 [505] 91 93 93 93 91 97 100 99 93 96 87 82 75 82 88 91 89 87  
 [523] 86 86 81 84 88 91 91 91 91 96 95 89 89 89 89 94 97 99  
 [541] 101 101 97 87 86 88 92 92 90 90 92 92 88 87 79 81 82 87  
 [559] 81 66 66 75 80 82 84 86 87 86 80 75 73 73 84 87 77 73  
 [577] 81 84 82 68 71 75 73 75 77 79 82 81 82 73 66 55 55 64  
 [595] 71 73 75 75 77 80 80 80 73 73 75 79 75 75 78 75 78 80  
 [613] 75 77 78 84 87 87 84 86 87 87 89 91 87 90 90 86 82 82  
 [631] 84 87 88 90 87 84 87 90 84 82 88 90 84 89 89 87 84 84  
 [649] 84 86 88 84 86 88 87 88 86 86 81 87 84 90 91 91 87 86  
 [667] 88 90 88 93 90 91 91 81 86 81 82 80 75 73 81 90 88 87  
 [685] 86 86 89 87 84 84 86 77 77 81 81 82 84 86 87 88 69 66  
 [703] 72 75 78 71 71 75 80 81 80 79 70 68 79 66 73 75 78 78  
 [721] 75 75 62 60 64 71 75 79 80 81 79 73 64 51 55 63 72 71  
 [739] 90 90 87 89 93 93 89 89 90 91 84 77 82 88 91 93 93 93  
 [757] 93 91 95 91 89 87 84 86 89 91 91 88 90 93 91 91 91 93  
 [775] 97 87 87 86 88 89 91 91 89 88 90 91 93 91 93 93 91 95  
 [793] 93 91 88 84 82 82 78 77 84 84 89 95 93 91 88 87 91 95  
 [811] 95 90 75 78 91 88 86 81 80 86 84 77 82 73 69 75 75 79  
 [829] 73 79 82 84 84 82 87 86 80 71 66 70 78 84 79 68 57 66  
 [847] 64 68 71 73 71 64 59 68 60 68 69 75 75 68 60 73 81 87  
 [865] 86 80 84 87 90 89 84 84 86 87 84 86 88 88 88 88 88 89  
 [883] 86 81 82 84 87 87 89 88 84 88 84 84 84 82 84 82 84 84  
 [901] 86 87 84 81 87 89 90 86 89 90 90 87 88 88 90 89 88 89  
 [919] 90 91 89 88 89 88 86 87 87 84 73 75 81 82 79 80 81 84  
 [937] 82 82 81 81 81 84 87 82 75 81 80 82 82 82 73 66 71 72  
 [955] 68 66 77 78 75 73 73 73 73 66 78 78 78 69 72 68 70 75  
 [973] 78 84 78 78 73 73 68 64 57 70 77 75 82 81 86 88 90 90  
 [991] 89 87 88 89 90 89 91 91 84 84  
 [ reached getOption("max.print") -- omitted 1460 entries ]

Hide

temps\_ts <- ts(temps\_mat, start=1996, end = 2015, frequency=123)  
temps\_ts

Time Series:  
Start = c(1996, 1)   
End = c(2015, 1)   
Frequency = 123   
 [1] 98 97 97 90 89 93 93 91 93 93 90 91 93 93 82 91 96 95  
 [19] 96 99 91 95 91 93 84 84 82 79 90 91 87 86 90 84 91 93  
 [37] 88 91 84 90 89 88 86 84 86 89 90 91 91 90 89 90 91 91  
 [55] 91 84 88 84 86 88 84 82 80 73 87 84 87 89 89 89 91 84  
 [73] 86 88 78 79 86 82 82 78 79 79 78 81 84 84 87 84 79 75  
 [91] 72 64 66 72 84 70 66 64 60 78 70 72 69 69 73 79 81 80  
 [109] 82 66 63 68 79 81 69 73 73 75 75 81 82 82 81 86 90 93  
 [127] 91 84 84 75 87 84 87 84 88 86 90 91 91 89 89 89 90 89  
 [145] 84 87 88 89 89 91 91 89 88 72 80 84 88 89 88 84 84 80  
 [163] 73 80 86 88 88 87 88 91 91 89 89 88 82 79 81 82 84 87  
 [181] 90 90 91 91 88 88 91 93 81 81 82 86 88 84 80 82 86 87  
 [199] 87 88 88 90 88 91 95 89 70 80 82 66 70 64 68 77 86 75  
 [217] 73 75 78 81 82 82 82 80 82 82 79 80 68 63 57 66 64 69  
 [235] 70 70 62 63 62 75 71 57 55 64 66 60 91 88 91 91 91 89  
 [253] 93 95 95 91 91 86 88 87 91 87 90 91 95 91 91 89 91 91  
 [271] 86 88 80 88 89 90 86 86 82 84 86 90 89 89 86 82 87 88  
 [289] 84 86 80 82 86 84 87 90 79 84 87 87 88 90 91 89 90 93  
 [307] 93 91 87 84 77 90 91 89 90 89 79 78 81 84 89 87 87 88  
 [325] 87 82 80 82 82 88 84 81 82 84 87 80 75 75 86 78 77 82  
 [343] 82 73 82 69 72 73 78 78 78 75 79 78 77 78 82 75 73 63  
 [361] 63 72 75 79 79 79 78 82 79 84 82 87 88 90 91 82 86 87  
 [379] 87 82 77 73 81 81 86 82 87 88 90 90 91 93 93 91 93 93  
 [397] 93 93 97 99 96 93 88 89 91 93 93 93 91 90 96 98 97 98  
 [415] 93 93 96 98 98 89 91 91 90 80 82 89 88 90 91 91 84 88  
 [433] 91 84 93 96 96 91 91 77 87 87 87 86 87 89 81 81 82 79  
 [451] 68 79 72 75 78 81 82 78 80 77 71 73 75 84 71 73 71 73  
 [469] 73 72 72 73 70 64 75 73 77 80 71 66 60 64 73 57 59 64  
 [487] 69 75 73 72 75 75 89 91 93 95 96 96 96 91 96 99 96 93  
 [505] 91 93 93 93 91 97 100 99 93 96 87 82 75 82 88 91 89 87  
 [523] 86 86 81 84 88 91 91 91 91 96 95 89 89 89 89 94 97 99  
 [541] 101 101 97 87 86 88 92 92 90 90 92 92 88 87 79 81 82 87  
 [559] 81 66 66 75 80 82 84 86 87 86 80 75 73 73 84 87 77 73  
 [577] 81 84 82 68 71 75 73 75 77 79 82 81 82 73 66 55 55 64  
 [595] 71 73 75 75 77 80 80 80 73 73 75 79 75 75 78 75 78 80  
 [613] 75 77 78 84 87 87 84 86 87 87 89 91 87 90 90 86 82 82  
 [631] 84 87 88 90 87 84 87 90 84 82 88 90 84 89 89 87 84 84  
 [649] 84 86 88 84 86 88 87 88 86 86 81 87 84 90 91 91 87 86  
 [667] 88 90 88 93 90 91 91 81 86 81 82 80 75 73 81 90 88 87  
 [685] 86 86 89 87 84 84 86 77 77 81 81 82 84 86 87 88 69 66  
 [703] 72 75 78 71 71 75 80 81 80 79 70 68 79 66 73 75 78 78  
 [721] 75 75 62 60 64 71 75 79 80 81 79 73 64 51 55 63 72 71  
 [739] 90 90 87 89 93 93 89 89 90 91 84 77 82 88 91 93 93 93  
 [757] 93 91 95 91 89 87 84 86 89 91 91 88 90 93 91 91 91 93  
 [775] 97 87 87 86 88 89 91 91 89 88 90 91 93 91 93 93 91 95  
 [793] 93 91 88 84 82 82 78 77 84 84 89 95 93 91 88 87 91 95  
 [811] 95 90 75 78 91 88 86 81 80 86 84 77 82 73 69 75 75 79  
 [829] 73 79 82 84 84 82 87 86 80 71 66 70 78 84 79 68 57 66  
 [847] 64 68 71 73 71 64 59 68 60 68 69 75 75 68 60 73 81 87  
 [865] 86 80 84 87 90 89 84 84 86 87 84 86 88 88 88 88 88 89  
 [883] 86 81 82 84 87 87 89 88 84 88 84 84 84 82 84 82 84 84  
 [901] 86 87 84 81 87 89 90 86 89 90 90 87 88 88 90 89 88 89  
 [919] 90 91 89 88 89 88 86 87 87 84 73 75 81 82 79 80 81 84  
 [937] 82 82 81 81 81 84 87 82 75 81 80 82 82 82 73 66 71 72  
 [955] 68 66 77 78 75 73 73 73 73 66 78 78 78 69 72 68 70 75  
 [973] 78 84 78 78 73 73 68 64 57 70 77 75 82 81 86 88 90 90  
 [991] 89 87 88 89 90 89 91 91 84 84  
 [ reached getOption("max.print") -- omitted 1338 entries ]

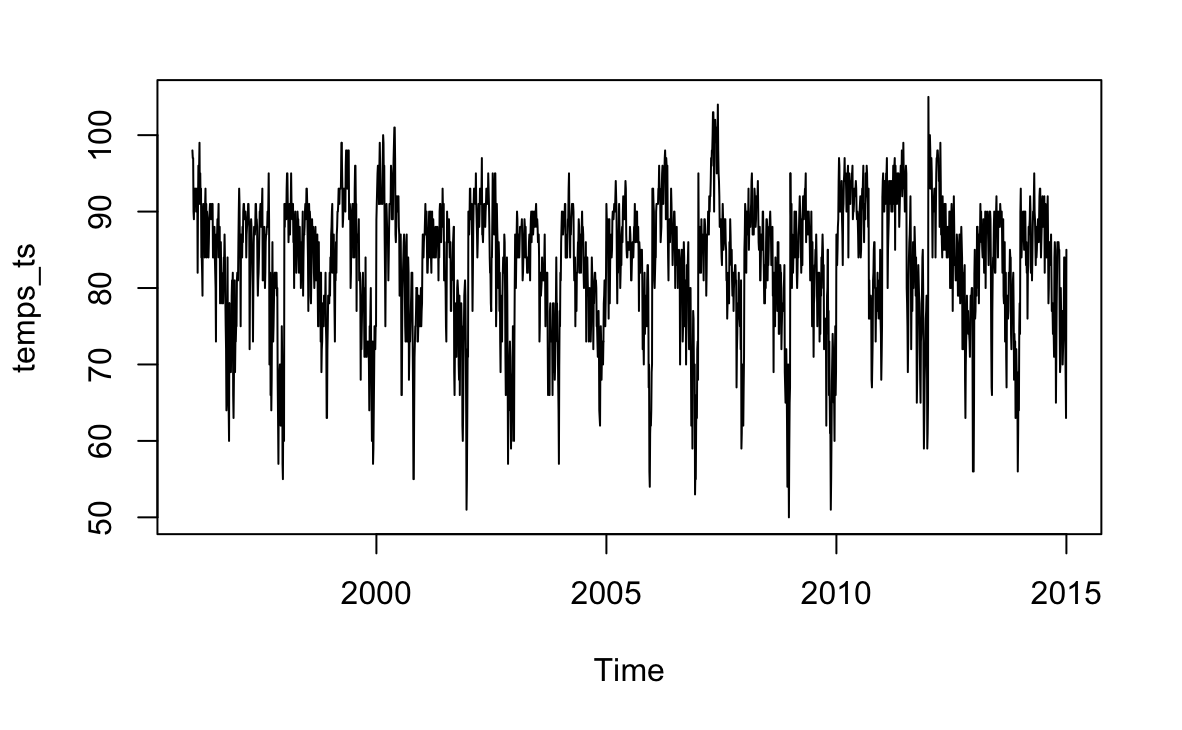
Hide

class(temps\_ts)

[1] "ts"

Hide

plot(temps\_ts)



Hide

#Exponential Smoothing   
#Simple Exponential   
temps\_single <- HoltWinters(temps\_ts,beta=FALSE, gamma=FALSE)  
#Double Exponential - model trend   
temps\_double <- HoltWinters(temps\_ts,gamma=FALSE)  
#Triple Exponential - model trend and seasonality  
temps\_triple\_additive <- HoltWinters(temps\_ts, seasonal = "additive")  
#Look at 3 kinds of ES  
temps\_single

Holt-Winters exponential smoothing without trend and without seasonal component.  
  
Call:  
HoltWinters(x = temps\_ts, beta = FALSE, gamma = FALSE)  
  
Smoothing parameters:  
 alpha: 0.8396301  
 beta : FALSE  
 gamma: FALSE  
  
Coefficients:  
 [,1]  
a 81.62444

Hide

temps\_single$SSE

[1] 53704.15

Hide

# Single ES parameters:  
# alpha: 0.8396301  
# SSE(sum of squared error):53704.15  
temps\_double

Holt-Winters exponential smoothing with trend and without seasonal component.  
  
Call:  
HoltWinters(x = temps\_ts, gamma = FALSE)  
  
Smoothing parameters:  
 alpha: 0.8455303  
 beta : 0.003777803  
 gamma: FALSE  
  
Coefficients:  
 [,1]  
a 81.729657393  
b -0.004838906

Hide

temps\_double$SSE

[1] 54071.22

Hide

# Double ES parameters:  
# alpha: 0.8455303  
# beta : 0.003777803  
# SSE: 54071.22  
temps\_triple

Holt-Winters exponential smoothing with trend and additive seasonal component.  
  
Call:  
HoltWinters(x = temps\_ts)  
  
Smoothing parameters:  
 alpha: 0.6677614  
 beta : 0  
 gamma: 0.6297674  
  
Coefficients:  
 [,1]  
a 66.739214602  
b -0.004362918  
s1 17.167113056  
s2 12.692593452  
s3 11.926233267  
s4 12.862822489  
s5 11.026083880  
s6 8.860499089  
s7 9.547553333  
s8 7.755384526  
s9 4.419013466  
s10 2.272689626  
s11 4.628251667  
s12 2.396834852  
s13 3.512957136  
s14 1.734948091  
s15 3.035023890  
s16 6.257944053  
s17 5.086362292  
s18 8.599153274  
s19 5.507486014  
s20 10.404819396  
s21 10.115801978  
s22 9.628840064  
s23 7.658623118  
s24 7.150473636  
s25 6.306599371  
s26 5.850691115  
s27 5.770487458  
s28 4.280481134  
s29 7.229771199  
s30 4.632381095  
s31 6.006248308  
s32 6.443645890  
s33 5.701166527  
s34 3.546887269  
s35 3.879569716  
s36 3.517339384  
s37 2.828550977  
s38 2.122971410  
s39 2.627923984  
s40 1.658896597  
s41 0.165866282  
s42 -0.001574460  
s43 -1.557500303  
s44 -2.159601227  
s45 -2.260609558  
s46 0.474052766  
s47 2.501631056  
s48 6.552191593  
s49 7.240238719  
s50 8.395899120  
s51 8.633263084  
s52 7.504540260  
s53 4.804135812  
s54 0.449902809  
s55 -1.045831475  
s56 1.562077049  
s57 1.632745190  
s58 0.857309158  
s59 2.909614779  
s60 0.626594899  
s61 4.491805650  
s62 4.567058619  
s63 3.065433531  
s64 3.787652805  
s65 -2.147135463  
s66 1.759895146  
s67 1.541155061  
s68 1.278521842  
s69 0.895959617  
s70 2.009912430  
s71 3.695537344  
s72 4.675235988  
s73 4.535880359  
s74 1.710420810  
s75 0.822675780  
s76 2.363162195  
s77 1.925012161  
s78 -1.656914701  
s79 -1.809929506  
s80 -0.427021203  
s81 0.056812125  
s82 -1.137248149  
s83 -1.037423821  
s84 -2.817503990  
s85 -4.578240308  
s86 -3.080091372  
s87 -2.710719111  
s88 -2.255335538  
s89 -4.518502545  
s90 -5.159556421  
s91 -4.440834373  
s92 -5.790113744  
s93 -7.461163074  
s94 -8.882612687  
s95 -8.619859733  
s96 -6.200719796  
s97 -6.055889182  
s98 -11.167287691  
s99 -13.489975101  
s100 -13.615536188  
s101 -14.373453486  
s102 -15.142110213  
s103 -14.419874185  
s104 -14.023613348  
s105 -16.187082843  
s106 -15.999259045  
s107 -12.074075053  
s108 -9.199729415  
s109 -10.403127076  
s110 -12.075113349  
s111 -9.722863134  
s112 -5.846856763  
s113 -8.047801338  
s114 -9.636669876  
s115 -10.510269852  
s116 -12.876648138  
s117 -8.657362442  
s118 -9.828539578  
s119 -14.522204766  
s120 -11.852457644  
s121 -8.714763993  
s122 -4.711332904  
s123 18.737998957

Hide

temps\_triple$SSE

[1] 63025.97

Hide

# Triple ES parameters:  
# alpha: 0.6677614  
# beta : 0  
# gamma: 0.6297674  
# SSE: 63025.97  
#Single ES gives the smallest SSE.Its alpha is closer to 1 which means there is less randomness in the system. The recent temperature observations have more weight in predicting the current temperature.  
#Seasonality can appear in two forms:   
#1. additive: amplitude of the seasonal variation is independent of the level,  
#2. multiplicative: amplitude of the seasonal variation is connected.   
#Triple Exponential - use multiplicative decomposition  
temps\_triple\_mul <- HoltWinters(temps\_ts, seasonal = "multiplicative")  
temps\_triple\_mul$SSE

[1] 65648.65

Hide

#SSE:65648.65  
#Triple Exponential - use additive decomposition  
temps\_triple\_additive <- HoltWinters(temps\_ts, seasonal = "additive")  
temps\_triple\_additive$SSE

[1] 63025.97

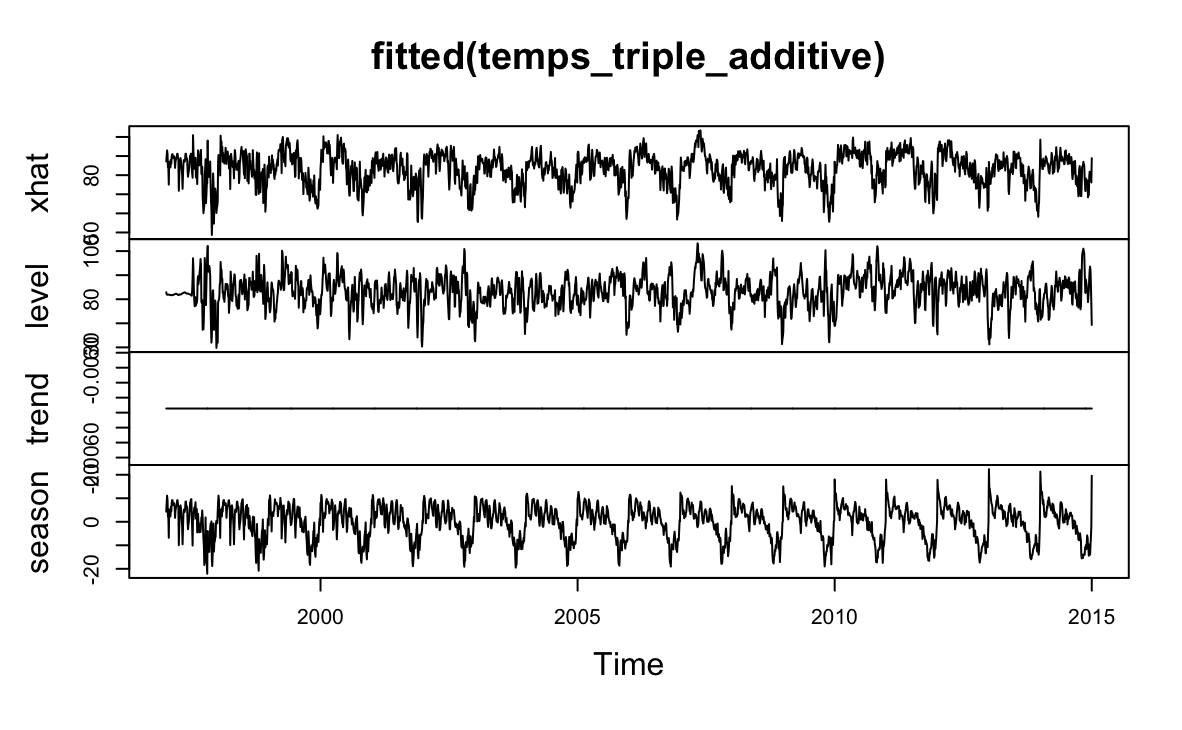
Hide

temps\_triple\_additive$fitted

Time Series:  
Start = c(1997, 1)   
End = c(2015, 1)   
Frequency = 123   
 xhat level trend season  
1997.000 87.17619 82.87739 -0.004362918 4.303159495  
1997.008 90.32137 82.08762 -0.004362918 8.238118845  
1997.016 92.95607 81.86865 -0.004362918 11.091777381  
1997.024 90.93226 81.89363 -0.004362918 9.042996893  
1997.033 83.99752 81.93450 -0.004362918 2.067387137  
1997.041 84.04359 81.93179 -0.004362918 2.116167625  
1997.049 75.06703 81.89832 -0.004362918 -6.826921806  
1997.057 87.04230 81.84919 -0.004362918 5.197468438  
1997.065 84.01782 81.81658 -0.004362918 2.205598519  
1997.073 87.05847 81.80032 -0.004362918 5.262509089  
1997.081 84.04758 81.75692 -0.004362918 2.295029414  
1997.089 88.04397 81.72078 -0.004362918 6.327549739  
1997.098 86.02650 81.68706 -0.004362918 4.343809902  
1997.106 89.93127 81.66500 -0.004362918 8.270639170  
1997.114 90.90776 81.70653 -0.004362918 9.205598519  
1997.122 90.94873 81.76376 -0.004362918 9.189338357  
1997.130 88.92982 81.79363 -0.004362918 7.140557869  
1997.138 88.90728 81.83613 -0.004362918 7.075517219  
1997.146 88.88353 81.89368 -0.004362918 6.994216406  
1997.154 89.85938 81.96709 -0.004362918 7.896655430  
1997.163 88.81884 82.05663 -0.004362918 6.766574129  
1997.171 83.84602 82.17324 -0.004362918 1.677143235  
1997.179 87.03391 82.27170 -0.004362918 4.766574129  
1997.187 88.03942 82.24469 -0.004362918 5.799094454  
1997.195 89.02500 82.21400 -0.004362918 6.815354617  
1997.203 89.17467 82.19295 -0.004362918 6.986086324  
1997.211 91.16749 82.07195 -0.004362918 9.099907462  
1997.220 91.17324 81.95574 -0.004362918 9.221858682  
1997.228 89.11010 81.83570 -0.004362918 7.278769251  
1997.236 87.99157 81.75781 -0.004362918 6.238118845  
1997.244 71.81397 81.75908 -0.004362918 -9.940742944  
1997.252 79.86066 81.87894 -0.004362918 -2.013913676  
1997.260 83.94121 81.96762 -0.004362918 1.977956243  
1997.268 88.04928 82.00251 -0.004362918 6.051126975  
1997.276 88.94697 81.96524 -0.004362918 6.986086324  
1997.285 87.85607 81.99629 -0.004362918 5.864135105  
1997.293 83.80148 82.08804 -0.004362918 1.717793641  
1997.301 83.75082 82.21625 -0.004362918 1.538931853  
1997.309 79.88033 82.37828 -0.004362918 -2.493588472  
1997.317 72.87458 82.45383 -0.004362918 -9.574889285  
1997.325 79.87267 82.53322 -0.004362918 -2.656190098  
1997.333 85.84764 82.61388 -0.004362918 3.238118845  
1997.341 87.86372 82.71126 -0.004362918 5.156818032  
1997.350 87.89345 82.79790 -0.004362918 5.099907462  
1997.358 87.04967 82.86469 -0.004362918 4.189338357  
1997.366 88.15848 82.82716 -0.004362918 5.335679820  
1997.374 91.23528 82.71697 -0.004362918 8.522671690  
1997.382 91.20389 82.55550 -0.004362918 8.652752991  
1997.390 89.07964 82.41499 -0.004362918 6.669013154  
1997.398 88.97332 82.35745 -0.004362918 6.620232666  
1997.407 87.97051 82.37090 -0.004362918 5.603972503  
1997.415 82.05901 82.38623 -0.004362918 -0.322856765  
1997.423 79.16971 82.34246 -0.004362918 -3.168385220  
1997.431 81.10080 82.22477 -0.004362918 -1.119604733  
1997.439 82.11856 82.15310 -0.004362918 -0.030173838  
1997.447 84.01877 82.06956 -0.004362918 1.953565999  
1997.455 87.03439 82.05267 -0.004362918 4.986086324  
1997.463 90.15341 82.02534 -0.004362918 8.132427788  
1997.472 90.25799 81.91854 -0.004362918 8.343809902  
1997.480 91.22769 81.74190 -0.004362918 9.490151365  
1997.488 91.20137 81.58550 -0.004362918 9.620232666  
1997.496 84.21295 81.44667 -0.004362918 2.770639170  
1997.504 80.81467 83.97116 -0.004362918 -3.152125058  
1997.512 78.66530 88.76488 -0.004362918 -10.095214489  
1997.520 100.93009 96.99715 -0.004362918 3.937305836  
1997.528 92.62219 91.69738 -0.004362918 0.929175755  
1997.537 87.89763 83.93216 -0.004362918 3.969826162  
1997.545 85.36047 79.32183 -0.004362918 6.042996893  
1997.553 83.25846 77.07348 -0.004362918 6.189338357  
1997.561 85.11731 78.89981 -0.004362918 6.221858682  
1997.569 89.11107 80.82040 -0.004362918 8.295029414  
1997.577 78.74251 77.40306 -0.004362918 1.343809902  
1997.585 81.62663 78.23840 -0.004362918 3.392590389  
1997.593 83.89598 78.48336 -0.004362918 5.416980633  
1997.602 75.35351 79.88398 -0.004362918 -4.526108798  
1997.610 84.15061 87.65670 -0.004362918 -3.501718554  
1997.618 92.97579 89.55504 -0.004362918 3.425110715  
1997.626 85.64879 86.22804 -0.004362918 -0.574889285  
1997.634 87.27138 87.79373 -0.004362918 -0.517978716  
1997.642 85.13787 89.61143 -0.004362918 -4.469198229  
1997.650 88.10164 91.51829 -0.004362918 -3.412287659  
1997.659 90.10586 93.44934 -0.004362918 -3.339116928  
1997.667 92.38587 96.71309 -0.004362918 -4.322856765  
1997.675 93.20999 94.44778 -0.004362918 -1.233425871  
1997.683 80.73941 78.94468 -0.004362918 1.799094454  
1997.691 80.28195 78.44657 -0.004362918 1.839744861  
1997.699 84.38418 79.58945 -0.004362918 4.799094454  
1997.707 69.06292 67.30884 -0.004362918 1.758444048  
1997.715 64.61113 67.93022 -0.004362918 -3.314726684  
1997.724 60.10112 67.51777 -0.004362918 -7.412287659  
1997.732 62.37945 72.78797 -0.004362918 -10.404157578  
1997.740 64.16252 82.54665 -0.004362918 -18.379767334  
1997.748 80.86233 97.12451 -0.004362918 -16.257816115  
1997.756 82.99211 93.20551 -0.004362918 -10.209035627  
1997.764 88.36418 86.52880 -0.004362918 1.839744861  
1997.772 65.40321 77.60035 -0.004362918 -12.192775464  
1997.780 69.82676 86.00764 -0.004362918 -16.176515302  
1997.789 75.32411 93.46433 -0.004362918 -18.135864895  
1997.797 75.81017 97.91788 -0.004362918 -22.103344570  
1997.805 97.99605 102.04684 -0.004362918 -4.046434001  
1997.813 79.34266 91.36094 -0.004362918 -12.013913676  
1997.821 81.91546 91.79552 -0.004362918 -9.875702294  
1997.829 79.04072 91.84761 -0.004362918 -12.802531562  
1997.837 81.02871 93.81934 -0.004362918 -12.786271399  
1997.846 83.65339 92.46029 -0.004362918 -8.802531562  
1997.854 87.17691 90.01633 -0.004362918 -2.835051887  
1997.862 76.35882 77.20636 -0.004362918 -0.843181968  
1997.870 66.44208 68.28150 -0.004362918 -1.835051887  
1997.878 62.12453 61.97208 -0.004362918 0.156818032  
1997.886 48.70806 64.55560 -0.004362918 -15.843181968  
1997.894 55.93132 74.76261 -0.004362918 -18.826921806  
1997.902 69.66185 83.48500 -0.004362918 -13.818791725  
1997.911 80.89142 83.70644 -0.004362918 -2.810661643  
1997.919 75.67923 76.42921 -0.004362918 -0.745620993  
1997.927 54.63797 67.29039 -0.004362918 -12.648060017  
1997.935 64.29875 72.86987 -0.004362918 -8.566759204  
1997.943 62.83254 71.33049 -0.004362918 -8.493588472  
1997.951 72.95314 79.45109 -0.004362918 -6.493588472  
1997.959 71.65267 78.14250 -0.004362918 -6.485458391  
1997.967 67.81504 68.35364 -0.004362918 -0.534238879  
1997.976 60.22077 59.79189 -0.004362918 0.433240796  
1997.984 62.71564 62.31115 -0.004362918 0.408850552  
1997.992 63.84754 64.49996 -0.004362918 -0.648060017  
1998.000 65.97906 61.92636 -0.004362918 4.057062406  
1998.008 86.79653 78.63002 -0.004362918 8.170877414  
1998.016 90.52589 79.42928 -0.004362918 11.100969520  
1998.024 88.79432 79.74151 -0.004362918 9.057170233  
1998.033 83.27356 81.21002 -0.004362918 2.067905689  
1998.041 88.46776 86.36507 -0.004362918 2.107046284  
1998.049 79.87081 86.71612 -0.004362918 -6.840946973  
1998.057 100.66318 95.47892 -0.004362918 5.188618358  
1998.065 93.89042 91.69291 -0.004362918 2.201869966  
1998.073 97.67540 92.42948 -0.004362918 5.250275614  
1998.081 90.24826 87.96755 -0.004362918 2.285073521  
1998.089 94.77916 88.46517 -0.004362918 6.318350547  
1998.098 86.93233 82.59842 -0.004362918 4.338264278  
1998.106 91.58767 83.30701 -0.004362918 8.285019302  
1998.114 89.45972 80.23918 -0.004362918 9.224897674  
1998.122 90.45906 81.26336 -0.004362918 9.200065311  
1998.130 86.10005 78.94917 -0.004362918 7.155241134  
1998.138 88.63960 81.54904 -0.004362918 7.094917086  
1998.146 90.13509 83.12087 -0.004362918 7.018585474  
1998.154 94.28682 86.36510 -0.004362918 7.926077622  
1998.163 90.96605 84.16593 -0.004362918 6.804479491  
1998.171 85.88924 84.18424 -0.004362918 1.709361624  
1998.179 91.01224 86.25712 -0.004362918 4.759479289  
1998.187 92.03107 86.24459 -0.004362918 5.790845815  
1998.195 92.35748 85.55172 -0.004362918 6.810124804  
1998.203 88.24725 81.30207 -0.004362918 6.949538929  
1998.211 90.19311 81.13261 -0.004362918 9.064862730  
1998.220 83.50293 74.32168 -0.004362918 9.185612110  
1998.228 84.57166 77.32029 -0.004362918 7.255732036  
1998.236 86.50852 80.27300 -0.004362918 6.239883268  
1998.244 72.69393 82.60011 -0.004362918 -9.901820071  
1998.252 89.49191 91.48103 -0.004362918 -1.984759393  
1998.260 91.13080 89.14491 -0.004362918 1.990256369  
1998.268 89.07980 83.04335 -0.004362918 6.040816683  
1998.276 86.63971 79.64689 -0.004362918 6.997182367  
1998.285 85.10524 79.21535 -0.004362918 5.894250731  
1998.293 84.23449 82.47952 -0.004362918 1.759331552  
1998.301 87.24409 85.65738 -0.004362918 1.591069014  
1998.309 84.35264 86.82555 -0.004362918 -2.468549302  
1998.317 78.36822 87.92123 -0.004362918 -9.548646623  
1998.325 87.70812 90.34203 -0.004362918 -2.629547590  
1998.333 93.13045 89.86481 -0.004362918 3.269997443  
1998.341 91.61551 86.43453 -0.004362918 5.185333016  
1998.350 86.46267 81.34483 -0.004362918 5.122201679  
1998.358 85.20610 81.03152 -0.004362918 4.178946324  
1998.366 82.84888 77.55072 -0.004362918 5.302520582  
1998.374 85.44859 76.97951 -0.004362918 8.473442899  
1998.382 85.94909 77.34336 -0.004362918 8.610092846  
1998.390 82.68546 76.03747 -0.004362918 6.652350510  
1998.398 85.53564 78.91419 -0.004362918 6.625816049  
1998.407 87.49673 81.89095 -0.004362918 5.610142549  
1998.415 75.87323 76.21280 -0.004362918 -0.335203707  
1998.423 78.42692 81.63518 -0.004362918 -3.203895037  
1998.431 86.21053 87.35559 -0.004362918 -1.140696088  
1998.439 87.81906 87.87840 -0.004362918 -0.054980233  
1998.447 89.94014 87.99486 -0.004362918 1.949639387  
1998.455 93.00500 88.03047 -0.004362918 4.978890292  
1998.463 94.78322 86.68725 -0.004362918 8.100330384  
1998.472 91.10654 82.82108 -0.004362918 8.289830648  
1998.480 91.51595 82.07781 -0.004362918 9.442510769  
1998.488 92.63817 83.06443 -0.004362918 9.578100181  
1998.496 86.86034 83.30169 -0.004362918 3.563016184  
1998.504 84.40855 86.06163 -0.004362918 -1.648715946  
1998.512 80.26899 87.78774 -0.004362918 -7.514388171  
1998.520 92.54851 90.27480 -0.004362918 2.278067894  
1998.528 78.38081 79.88775 -0.004362918 -1.502570622  
1998.537 90.16448 87.64223 -0.004362918 2.526614473  
1998.545 93.53131 88.19579 -0.004362918 5.339876352  
1998.553 91.92420 85.16560 -0.004362918 6.762960169  
1998.561 90.69698 83.87633 -0.004362918 6.825012625  
1998.569 89.96005 82.73879 -0.004362918 7.225625268  
1998.577 77.01828 75.41573 -0.004362918 1.606918095  
1998.585 79.53327 76.06692 -0.004362918 3.470711462  
1998.593 82.89483 77.04198 -0.004362918 5.857211189  
1998.602 75.68197 77.77561 -0.004362918 -2.089278380  
1998.610 83.75462 86.66451 -0.004362918 -2.905532868  
1998.618 91.20694 88.82729 -0.004362918 2.384010915  
1998.626 85.92640 86.01370 -0.004362918 -0.082937756  
1998.634 87.44258 87.39401 -0.004362918 0.052937961  
1998.642 83.21940 87.09410 -0.004362918 -3.870345199  
1998.650 83.46526 86.27548 -0.004362918 -2.805854177  
1998.659 81.63768 83.95715 -0.004362918 -2.315100992  
1998.667 79.15907 84.19473 -0.004362918 -5.031293693  
1998.675 79.99334 86.08743 -0.004362918 -6.089722075  
1998.683 93.06962 91.42960 -0.004362918 1.644385312  
1998.691 87.56375 85.36889 -0.004362918 2.199218069  
1998.699 81.92967 80.98151 -0.004362918 0.952515668  
1998.707 82.97427 81.02412 -0.004362918 1.954511737  
1998.715 78.25774 81.70470 -0.004362918 -3.442595878  
1998.724 81.77414 87.53808 -0.004362918 -5.759580500  
1998.732 78.99960 86.34902 -0.004362918 -7.345056098  
1998.740 69.85887 83.67388 -0.004362918 -13.810644299  
1998.748 69.61379 87.10256 -0.004362918 -17.484410097  
1998.756 85.73620 98.04028 -0.004362918 -12.299715505  
1998.764 91.90913 92.86998 -0.004362918 -0.956483800  
1998.772 73.34840 82.90987 -0.004362918 -9.557111515  
1998.780 74.83965 88.68272 -0.004362918 -13.838703967  
1998.789 76.71635 93.45976 -0.004362918 -16.739047137  
1998.797 70.16117 90.97376 -0.004362918 -20.808227064  
1998.805 91.47721 98.87491 -0.004362918 -7.393335572  
1998.813 71.98039 83.86113 -0.004362918 -11.876376715  
1998.821 74.00749 83.86986 -0.004362918 -9.858012730  
1998.829 71.00502 83.19274 -0.004362918 -12.183352127  
1998.837 74.64424 87.85935 -0.004362918 -13.210744539  
1998.846 80.52453 90.09583 -0.004362918 -9.566942039  
1998.854 81.55383 88.40569 -0.004362918 -6.847496242  
1998.862 80.38228 84.02493 -0.004362918 -3.638288239  
1998.870 79.28252 83.09753 -0.004362918 -3.810647528  
1998.878 83.20008 82.23675 -0.004362918 0.967694205  
1998.886 65.44425 78.09221 -0.004362918 -12.643602052  
1998.894 70.37521 86.47210 -0.004362918 -16.092522774  
1998.902 80.47792 94.23032 -0.004362918 -13.748039245  
1998.911 85.47415 90.56801 -0.004362918 -5.089506695  
1998.919 78.62177 82.23390 -0.004362918 -3.607767550  
1998.927 60.89511 71.79792 -0.004362918 -10.898446041  
1998.935 64.14703 73.19912 -0.004362918 -9.047733112  
1998.943 72.48656 78.43867 -0.004362918 -5.947753797  
1998.951 73.20608 80.11269 -0.004362918 -6.902249366  
1998.959 74.42164 83.97728 -0.004362918 -9.551282217  
1998.967 83.81024 87.03017 -0.004362918 -3.215568915  
1998.976 85.03334 83.81372 -0.004362918 1.223981060  
1998.984 80.20445 79.11276 -0.004362918 1.096047196  
1998.992 78.84995 80.30740 -0.004362918 -1.453091577  
1999.000 89.69115 80.40324 -0.004362918 9.292270132  
1999.008 85.01687 76.59855 -0.004362918 8.422683133  
1999.016 85.77544 74.57964 -0.004362918 11.200169006  
1999.024 84.90729 75.39299 -0.004362918 9.518671355  
 [ reached getOption("max.print") -- omitted 1965 rows ]

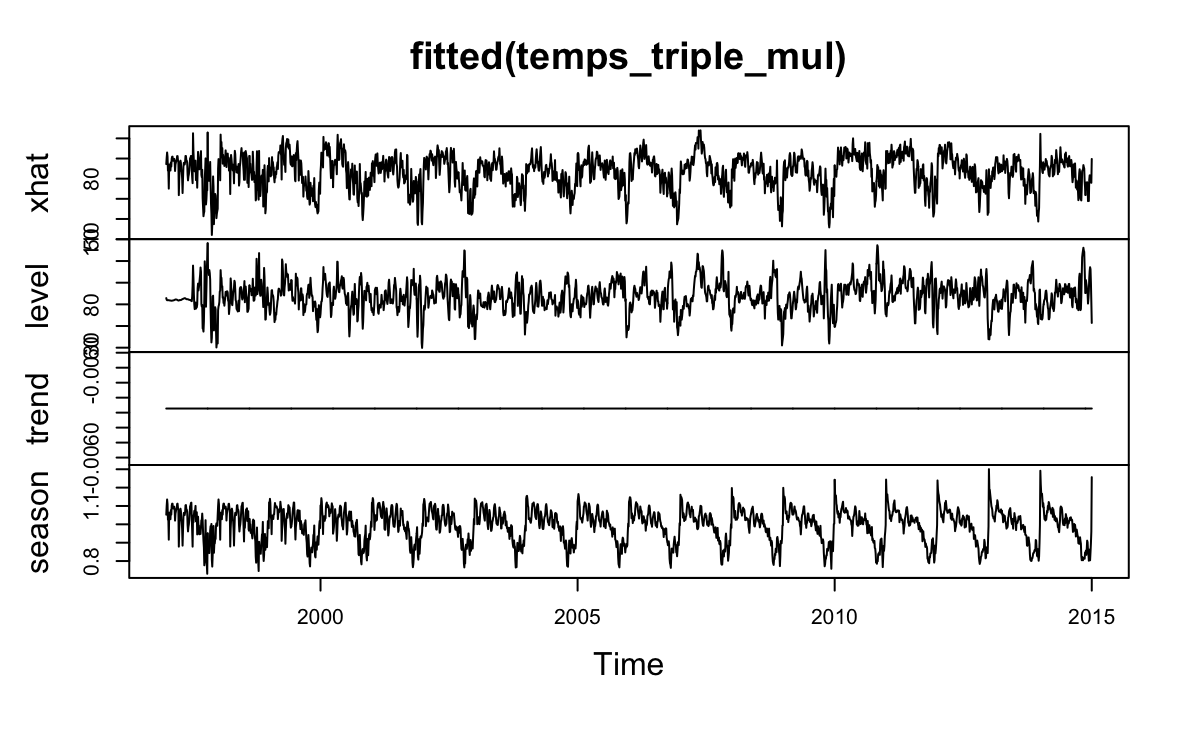
Hide

#SSE:63025.97  
plot(fitted(temps\_triple\_additive))



Hide

plot(fitted(temps\_triple\_mul))



Triple ES with additive seasonal factor has better SSE.

From looking at the chart: the trend subchart shows a straight line which means there is no trend; this is also evidenced by alpha=0.

The season subchart shows that thee duration of each season has been pretty constant throughout these years. The next step I would do is to apply cumsum method to the level data for the daily temperature from 1997 to 2005 for each year and set the C and T values to see whether the last day of summer has become earlier or later.

Question 8.1 Describe a situation or problem from your job, everyday life, current events, etc., for which a linear regression model would be appropriate. List some (up to 5) predictors that you might use.

The credit card company will analyze cardholder information to predict the balance of a new cardholder for customer analysis purpose. Some predictors to be used are: 1. the past monthly balance 2. the cardholder’s income 3. the industry cardholder works in 4. spending on essential goods (food, medical expenses) 5. spending on nonessential goods (luxury brands, travelling, etc.)

Question 8.2 Using crime data from <http://www.statsci.org/data/general/uscrime.txt> (file uscrime.txt, description at <http://www.statsci.org/data/general/uscrime.html> ), use regression (a useful R function is lm or glm) to predict the observed crime rate in a city with the following data: M = 14.0 So = 0 Ed = 10.0 Po1 = 12.0 Po2 = 15.5 LF = 0.640 M.F = 94.0 Pop = 150 NW = 1.1 U1 = 0.120 U2 = 3.6 Wealth = 3200 Ineq = 20.1 Prob = 0.04 Time = 39.0

Show your model (factors used and their coefficients), the software output, and the quality of fit.

Note that because there are only 47 data points and 15 predictors, you’ll probably notice some overfitting. We’ll see ways of dealing with this sort of problem later in the course.

Hide

Data\_crime = read.csv("uscrime.txt",sep = "")  
str(Data\_crime)

'data.frame': 47 obs. of 16 variables:  
 $ M : num 15.1 14.3 14.2 13.6 14.1 12.1 12.7 13.1 15.7 14 ...  
 $ So : int 1 0 1 0 0 0 1 1 1 0 ...  
 $ Ed : num 9.1 11.3 8.9 12.1 12.1 11 11.1 10.9 9 11.8 ...  
 $ Po1 : num 5.8 10.3 4.5 14.9 10.9 11.8 8.2 11.5 6.5 7.1 ...  
 $ Po2 : num 5.6 9.5 4.4 14.1 10.1 11.5 7.9 10.9 6.2 6.8 ...  
 $ LF : num 0.51 0.583 0.533 0.577 0.591 0.547 0.519 0.542 0.553 0.632 ...  
 $ M.F : num 95 101.2 96.9 99.4 98.5 ...  
 $ Pop : int 33 13 18 157 18 25 4 50 39 7 ...  
 $ NW : num 30.1 10.2 21.9 8 3 4.4 13.9 17.9 28.6 1.5 ...  
 $ U1 : num 0.108 0.096 0.094 0.102 0.091 0.084 0.097 0.079 0.081 0.1 ...  
 $ U2 : num 4.1 3.6 3.3 3.9 2 2.9 3.8 3.5 2.8 2.4 ...  
 $ Wealth: int 3940 5570 3180 6730 5780 6890 6200 4720 4210 5260 ...  
 $ Ineq : num 26.1 19.4 25 16.7 17.4 12.6 16.8 20.6 23.9 17.4 ...  
 $ Prob : num 0.0846 0.0296 0.0834 0.0158 0.0414 ...  
 $ Time : num 26.2 25.3 24.3 29.9 21.3 ...  
 $ Crime : int 791 1635 578 1969 1234 682 963 1555 856 705 ...

Hide

print(summary(Data\_crime))

M So Ed Po1   
 Min. :11.90 Min. :0.0000 Min. : 8.70 Min. : 4.50   
 1st Qu.:13.00 1st Qu.:0.0000 1st Qu.: 9.75 1st Qu.: 6.25   
 Median :13.60 Median :0.0000 Median :10.80 Median : 7.80   
 Mean :13.86 Mean :0.3404 Mean :10.56 Mean : 8.50   
 3rd Qu.:14.60 3rd Qu.:1.0000 3rd Qu.:11.45 3rd Qu.:10.45   
 Max. :17.70 Max. :1.0000 Max. :12.20 Max. :16.60   
 Po2 LF M.F Pop   
 Min. : 4.100 Min. :0.4800 Min. : 93.40 Min. : 3.00   
 1st Qu.: 5.850 1st Qu.:0.5305 1st Qu.: 96.45 1st Qu.: 10.00   
 Median : 7.300 Median :0.5600 Median : 97.70 Median : 25.00   
 Mean : 8.023 Mean :0.5612 Mean : 98.30 Mean : 36.62   
 3rd Qu.: 9.700 3rd Qu.:0.5930 3rd Qu.: 99.20 3rd Qu.: 41.50   
 Max. :15.700 Max. :0.6410 Max. :107.10 Max. :168.00   
 NW U1 U2 Wealth   
 Min. : 0.20 Min. :0.07000 Min. :2.000 Min. :2880   
 1st Qu.: 2.40 1st Qu.:0.08050 1st Qu.:2.750 1st Qu.:4595   
 Median : 7.60 Median :0.09200 Median :3.400 Median :5370   
 Mean :10.11 Mean :0.09547 Mean :3.398 Mean :5254   
 3rd Qu.:13.25 3rd Qu.:0.10400 3rd Qu.:3.850 3rd Qu.:5915   
 Max. :42.30 Max. :0.14200 Max. :5.800 Max. :6890   
 Ineq Prob Time Crime   
 Min. :12.60 Min. :0.00690 Min. :12.20 Min. : 342.0   
 1st Qu.:16.55 1st Qu.:0.03270 1st Qu.:21.60 1st Qu.: 658.5   
 Median :17.60 Median :0.04210 Median :25.80 Median : 831.0   
 Mean :19.40 Mean :0.04709 Mean :26.60 Mean : 905.1   
 3rd Qu.:22.75 3rd Qu.:0.05445 3rd Qu.:30.45 3rd Qu.:1057.5   
 Max. :27.60 Max. :0.11980 Max. :44.00 Max. :1993.0

Hide

model <- lm(Crime ~ . ,Data\_crime)  
model

Call:  
lm(formula = Crime ~ ., data = Data\_crime)  
  
Coefficients:  
(Intercept) M So Ed Po1 Po2   
 -5.984e+03 8.783e+01 -3.803e+00 1.883e+02 1.928e+02 -1.094e+02   
 LF M.F Pop NW U1 U2   
 -6.638e+02 1.741e+01 -7.330e-01 4.204e+00 -5.827e+03 1.678e+02   
 Wealth Ineq Prob Time   
 9.617e-02 7.067e+01 -4.855e+03 -3.479e+00

Hide

print(summary(model))

Call:  
lm(formula = Crime ~ ., data = Data\_crime)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-395.74 -98.09 -6.69 112.99 512.67   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) -5.984e+03 1.628e+03 -3.675 0.000893 \*\*\*  
M 8.783e+01 4.171e+01 2.106 0.043443 \*   
So -3.803e+00 1.488e+02 -0.026 0.979765   
Ed 1.883e+02 6.209e+01 3.033 0.004861 \*\*   
Po1 1.928e+02 1.061e+02 1.817 0.078892 .   
Po2 -1.094e+02 1.175e+02 -0.931 0.358830   
LF -6.638e+02 1.470e+03 -0.452 0.654654   
M.F 1.741e+01 2.035e+01 0.855 0.398995   
Pop -7.330e-01 1.290e+00 -0.568 0.573845   
NW 4.204e+00 6.481e+00 0.649 0.521279   
U1 -5.827e+03 4.210e+03 -1.384 0.176238   
U2 1.678e+02 8.234e+01 2.038 0.050161 .   
Wealth 9.617e-02 1.037e-01 0.928 0.360754   
Ineq 7.067e+01 2.272e+01 3.111 0.003983 \*\*   
Prob -4.855e+03 2.272e+03 -2.137 0.040627 \*   
Time -3.479e+00 7.165e+00 -0.486 0.630708   
---  
Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
Residual standard error: 209.1 on 31 degrees of freedom  
Multiple R-squared: 0.8031, Adjusted R-squared: 0.7078   
F-statistic: 8.429 on 15 and 31 DF, p-value: 3.539e-07

I apply all the predictors as my independent variables in the linear regression model. The adjusted R-Square (quality of fit) of 0.7 is not bad. The p-value for each predictor tests the null hypothesis that the coefficient is equal to zero (no effect). A low p-value (< 0.05) indicates that you can reject the null hypothesis.

Given there are too many predictors for too few number of observations (probably overfitting issue), I tried to remove some predictors with high p-values which are: So, Po2, LF, M.F, Pop, NW, U1,Wealth,Time. I create this new lr model below:

Hide

model\_2 <- lm(Crime ~ M+Ed+Po1+U2+Ineq+Prob ,Data\_crime)  
print(summary(model\_2))

Call:  
lm(formula = Crime ~ M + Ed + Po1 + U2 + Ineq + Prob, data = Data\_crime)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-470.68 -78.41 -19.68 133.12 556.23   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) -5040.50 899.84 -5.602 1.72e-06 \*\*\*  
M 105.02 33.30 3.154 0.00305 \*\*   
Ed 196.47 44.75 4.390 8.07e-05 \*\*\*  
Po1 115.02 13.75 8.363 2.56e-10 \*\*\*  
U2 89.37 40.91 2.185 0.03483 \*   
Ineq 67.65 13.94 4.855 1.88e-05 \*\*\*  
Prob -3801.84 1528.10 -2.488 0.01711 \*   
---  
Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
Residual standard error: 200.7 on 40 degrees of freedom  
Multiple R-squared: 0.7659, Adjusted R-squared: 0.7307   
F-statistic: 21.81 on 6 and 40 DF, p-value: 3.418e-11

Interestingly, removing some predictors actually help improve the model’s adjusted R-squared.

Hide

# Use the first linear regression model to make the prediction  
test<-data.frame(M = 14.0,So = 0,Ed = 10.0, Po1 = 12.0,Po2 = 15.5,  
 LF = 0.640, M.F = 94.0,Pop = 150,NW = 1.1,U1 = 0.120,  
 U2 = 3.6, Wealth = 3200,Ineq = 20.1,Prob = 0.04, Time = 39.0)  
print(predict(model,test))

1   
155.4349