Assignment 3

IM 532 3.0 Applied Time Series Forecasting

14/06/2020

Question 1

Data set `fma::books` contains the daily sales of paperback and hardcover books at the same store.

```
library(fma)
books
```

```
Time Series:
Start = 1
End = 30
Frequency = 1
   Paperback Hardcover
 1       199       139
 2       172       128
 3       111       172
 4       209       139
 5       161       191
 6       119       168
 7       195       170
 8       195       145
 9       131       184
10       183       135
11       143       218
12       141       198
13       168       230
14       201       222
15       155       206
16       243       240
17       225       189
18       167       222
19       237       158
20       202       178
21       186       217
22       176       261
23       232       238
24       195       240
25       190       214
26       182       200
27       222       201
28       217       283
29       188       220
30       247       259
```

The task is to forecast the next four days' sales for paperback and hardcover books.

i) Plot the data and describe the main features of the series.

ii) Use an ETS(A,N,N) model to forecast each series, and plot the forecasts.

iii) Compute the RMSE values for the training data in each case.

Question 2

For this exercise use data set `expsmooth::visitors`, the monthly Australian short-term overseas visitors data, May 1985–April 2005

```
library(expsmooth)
visitors
```

```
       Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec
1985                           75.7  75.4  83.1  82.9  77.3 105.7 121.9 150.0
1986  98.0 118.0 129.5 110.6  91.7  94.8 109.5 105.1  95.0 130.3 156.7 190.1
1987 139.7 147.8 145.2 132.7 120.7 116.5 142.0 140.4 128.0 165.7 183.1 222.8
1988 161.3 180.4 185.2 160.5 157.1 163.8 203.3 196.9 179.6 207.3 208.0 245.8
1989 168.9 191.1 180.0 160.1 136.6 142.7 175.4 161.4 149.9 174.1 192.7 247.4
1990 176.2 192.8 189.1 181.1 149.9 157.3 185.3 178.2 162.7 190.6 198.6 253.1
1991 177.4 190.6 189.2 168.0 161.4 172.2 208.3 199.3 197.4 216.0 223.9 266.8
1992 196.1 238.2 217.8 203.8 175.2 176.9 219.3 199.1 190.0 229.3 255.0 302.4
1993 242.8 245.5 257.9 226.3 213.4 204.6 244.6 239.9 224.0 267.2 285.9 344.0
1994 250.5 304.3 307.4 255.1 214.9 230.9 282.5 265.4 254.0 301.6 311.0 384.0
1995 303.8 319.1 313.5 294.2 244.8 261.4 329.7 304.9 268.6 320.7 342.9 422.3
1996 317.2 392.7 365.6 333.2 261.5 306.9 358.2 329.2 309.2 350.4 375.6 465.2
1997 342.9 408.0 390.9 325.9 289.1 308.2 397.4 330.4 330.9 366.5 379.5 448.3
1998 346.2 353.6 338.6 341.1 283.4 304.2 372.3 323.7 323.9 354.8 367.9 457.6
1999 351.0 398.6 389.0 334.1 298.1 317.1 388.5 355.6 353.1 397.0 416.7 460.8
2000 360.8 434.6 411.9 405.6 319.3 347.9 429.0 372.9 403.0 426.5 459.9 559.9
2001 416.6 429.2 428.7 405.4 330.2 370.0 446.9 384.6 366.3 378.5 376.2 523.2
2002 379.3 437.2 446.5 360.3 329.9 339.4 418.2 371.9 358.6 428.9 437.0 534.0
2003 396.6 427.5 392.5 321.5 260.9 308.3 415.5 362.2 385.6 435.3 473.3 566.6
2004 420.2 454.8 432.3 402.8 341.3 367.3 472.0 405.8 395.6 449.9 479.9 593.1
2005 462.4 501.6 504.7 409.5
```

i. Make a time plot of your data and describe the main features of the series.

ii. Create a training set that withholds the last two years of available data.

iii. Use ARIMA, snaive, random walk, random walk with drift, average method on the time series. Which method gives the best forecasts? Give reasons for your answer.

Due date: 20 June 2020