**Assignment 2**

**Author:** Keerthy Raaj Shanmugam, 8779954

**Course:** PROG8430

**SECTION 1: Woodstock Temperature**

* 1. **Read in the Woodstock data and transform it into an appropriate time series datatype.**

Temp

[1,] -14.318013

[2,] -10.189734

[3,] -5.447450

[4,] 4.505676

[5,] 6.889375

[6,] 12.112377

The above-given output is the first 6 rows of Woodstock data which is converted into appropriate time series datatype.

**2.1 Summarize the temperature information (mean, etc.)**

Temp

Min. :-14.318

1st Qu.: -1.474

Median : 5.311

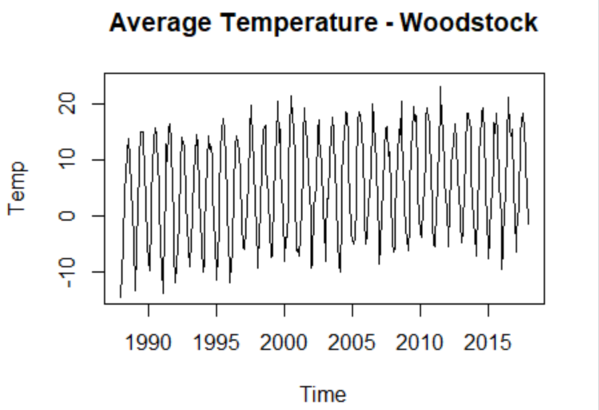
Mean : 5.654

3rd Qu.: 13.869

Max. : 23.118

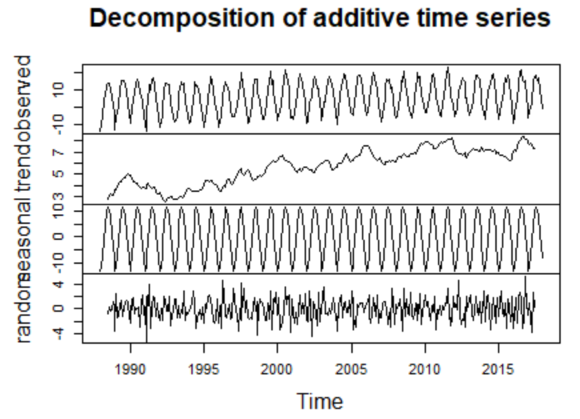
The above-given output is the summarized output information.

**2.2 Plot the time series data.**



The above-given chart is the timeseries of Woodstock temperature data.

**2.3 Decompose the times series data in to the constituent components. Comment on each (any trends you observe, etc.)**

****

**Random -** The fluctuations of temperature are random here.

**Seasonal -** There is a steady increase and decrease in fluctuation of temperature on alternate years.

**Trend -** There is minor fluctuations of temperature in alternate years but increases steadily till the end.

**Observed -** There is a steady increase and decrease in fluctuation of temperature on alternate years.

**2.4 Determine if the time series is stationary.**

Augmented Dickey-Fuller Test

data: Temp\_TS\_KS

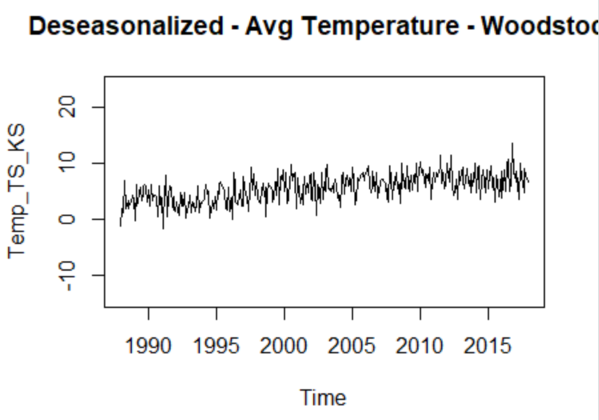
Dickey-Fuller = -13.173, Lag order = 7, p-value

= 0.01

alternative hypothesis: stationary

The p-value is 0.01 and it is less than 0.05, so the time series is **stationary.**

**2.5 Deseasonalize the information and plot the result.**

****

The above-given chart has temperature fluctuation between -2 to 12.

**2.6 Add any comments about what you observe: seasonality of temperature, trends, etc.**

**Random -** The fluctuations of temperature trend are random here.

**Seasonal -** There is a steady increase and decrease in fluctuation of temperature trend on alternate years.

**Trend -** There is minor fluctuations of temperature trend in alternate years but increases steadily till the end.

**Observed -** There is a steady increase and decrease in fluctuation of temperature trend on alternate years.

**SECTION 2: Ayr Temperature**

* 1. **Read in the Ayr data and transform it into an appropriate time series datatype.**

Temp

[1,] 11.20966

[2,] 12.85350

[3,] 12.10332

[4,] 10.98670

[5,] 12.78706

[6,] 12.47640

The above-given output is the first 6 rows of Ayr data which is converted into appropriate time series datatype.

**2.1 Summarize the information (mean, std dev, etc.)**

Temp

Min. :10.99

1st Qu.:12.27

Median :12.72

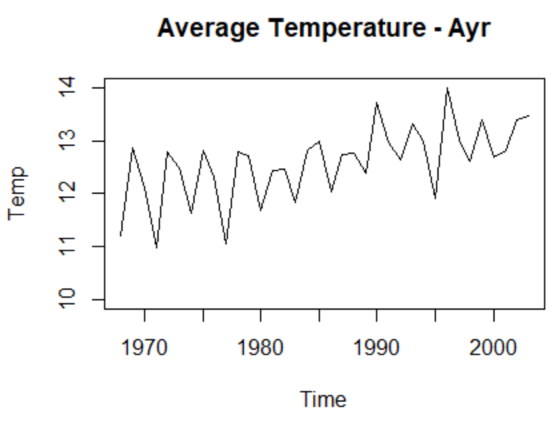
Mean :12.57

3rd Qu.:12.96

Max. :13.98

The above-given output is the summarized output information.

**2.2 Plot the time series data.**

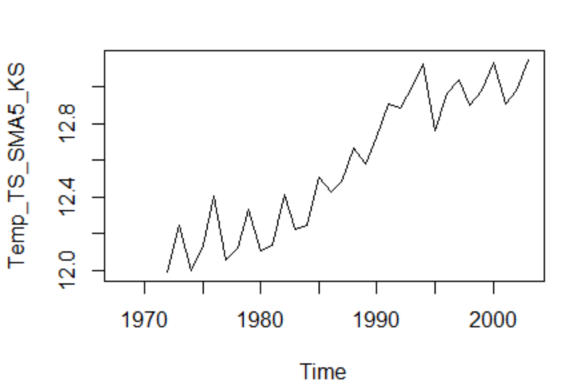
****

The above-given chart is the timeseries of Ayr temperature data.

**2.3 Smooth the temperature chart using a moving average. Try 3 different values for the moving average and choose the one you think best shows the trend (if any).**

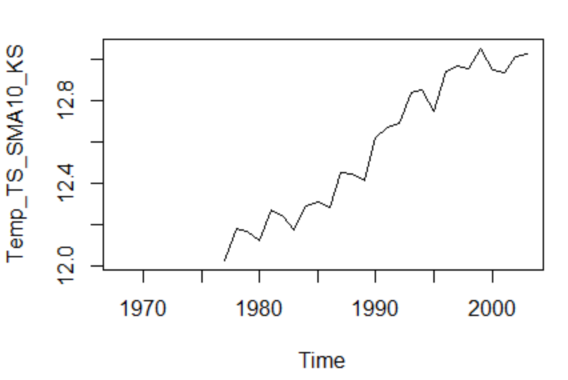
Temp\_TS\_SMA5\_KS <-SMA(Temp\_TS\_KS,n=5)

plot.ts(Temp\_TS\_SMA5\_KS)

****

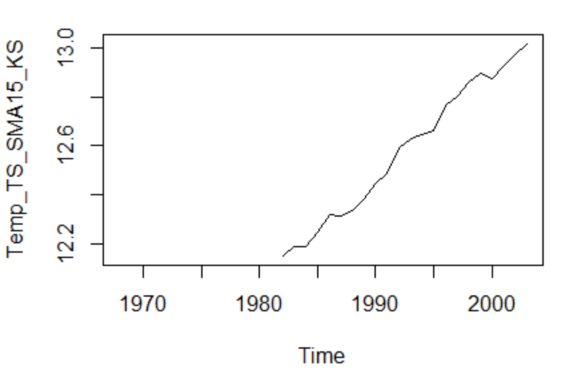
Temp\_TS\_SMA10\_KS <-SMA(Temp\_TS\_KS,n=10)

plot.ts(Temp\_TS\_SMA10\_KS)



Temp\_TS\_SMA15\_KS <-SMA(Temp\_TS\_KS,n=15)

plot.ts(Temp\_TS\_SMA15\_KS)



I used three values for the moving average they are **5,10 and 15**.

The best one to show the trend is when moving average = 10 because the trends are less craggy and smooth. Moreover, trend gives sufficient data, not too much or too less.

**2.4 Determine if the time series is stationary.**

Augmented Dickey-Fuller Test

data: Temp\_TS\_KS

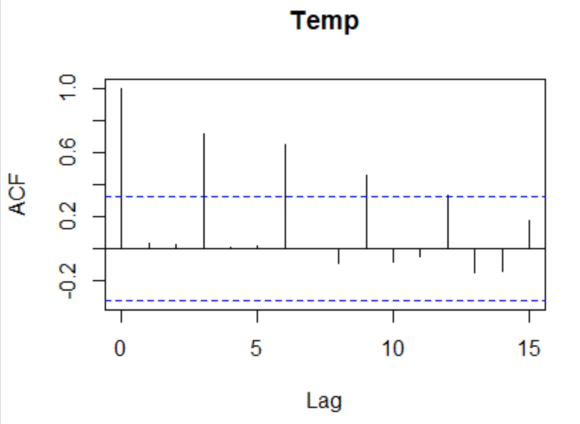
Dickey-Fuller = -3.2533, Lag order = 3, p-value

= 0.09458

alternative hypothesis: stationary

The p-value is 0.09458 and it is greater than 0.05, so the time series may be **non** **stationary.**

**2.5 Create an autocorrelation chart (using acf) and comment on which lags are significant. Do previous values seem to influence current values?**

****

Every vertical line is the correlation between current and previous observation. There are 3 lags that are significant and they are **lag 3, lag 6 and lag 9.**

**3.1 Create a simple moving average forecast of temperature in Ayr for five years beyond the data provided. Graph your results along with a 75% prediction interval.**

Time Series:

Start = 2004

End = 2008

Frequency = 1

Point forecast Lower bound (12.5%)

2004 13.06027 12.35296

2005 13.13431 12.41724

2006 13.09171 12.36157

2007 13.16109 12.41353

2008 13.21785 12.44721

Upper bound (87.5%)

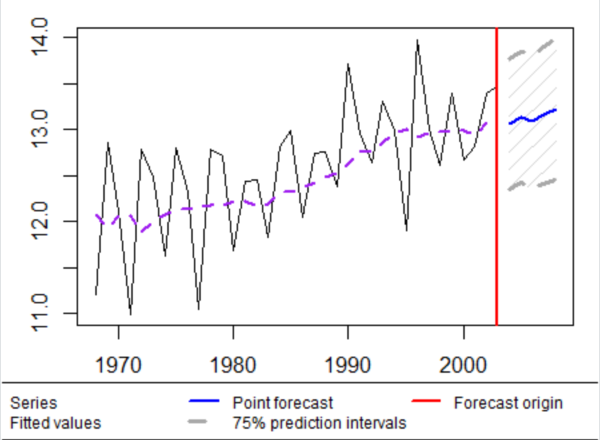
2004 13.76758

2005 13.85137

2006 13.82185

2007 13.90865

2008 13.98849



**3.2 Create an exponentially smoothed forecast of temperature in Ayr for five years beyond the data provided. Graph your results along with a 75% prediction interval.**

Time Series:

Start = 2004

End = 2008

Frequency = 1

Point forecast Lower bound (12.5%)

2004 13.30069 12.64099

2005 13.33993 12.68023

2006 13.37917 12.71947

2007 13.41842 12.75871

2008 13.45766 12.79795

Upper bound (87.5%)

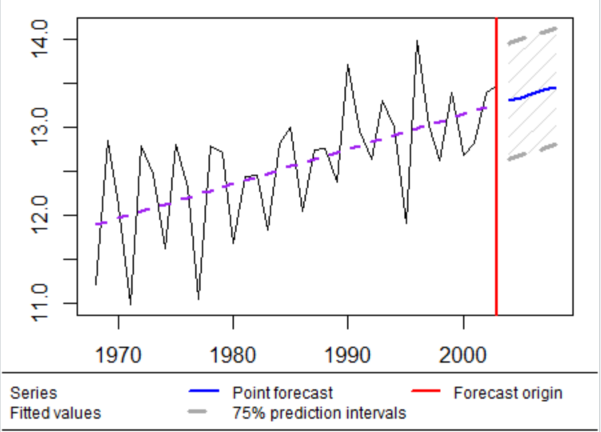
2004 13.96039

2005 13.99964

2006 14.03888

2007 14.07812

2008 14.11736



**3.3 Compare the two forecasts you created in steps 1 and 2 above. Which forecast seems superior? Why?**

**Step 1 - moving average forecast**

**Step 2 - exponentially smoothed forecast**

**Exponentially smoothed forecast** seems superior than **moving average forecast** becauseexponentially smoothed forecast had upward trend and it continuous to maintain the upward trend in forecasts.

**4.1 Your submission will be evaluated on the basis of professionalism and clarity of presentation.**