Assignment 1

Vinod Kumar Ahuja

September 21, 2017

Introduction

In this assignment, the data of GISS Surface Temperature Analysis (GISTEMP) is analyzed to check whether global temperature is rising over the period of time or not.

Data for the analysis is for the period from 1880 to 2015.

Detailed analysis is given below:

Data details:

First of all, data is loaded into R so that some visualization can be performed on this Data. Since this data contains time element so data is converted into time series format for R so that visualization can be performed accordingly.

Data provided was in two files.

1. The first monthly temperature of the globe from the periods 1880 to 2015.Besides monthly temperature, this data also contains quarterly data for the same period.
2. The temperature in different global zones from the north pole to south pole is a period for the same period i.e.1880 to 2015.

This data is analyzed in three different steps as follows.

1. In the first step, data is extracted only for all the months in from the entire data set to analyze it in time series format.
2. In a second step, quarterly data is analyzed.
3. In the third step, global zone data distributed from north pole to south pole is analyzed.

Detailed Analysis of each step is given below:

Step 1 (Analysis of Monthly Data)

First of all monthly data for the entire set from the year 1880 to 2014 is extracted in separate CSV files. The year 2015 is dropped as it does not contain data for all the months. This data have been loaded into R as shown in following steps:

```
# Reading Raw Data1
TData1<-scan("E:/onedrive/University of Nebraska Omaha/Courses/Data Visualization/Assignment 1/NASA GIS
```

Loaded couple of libraries for working with data.

```
library("TTR") # for SMA function
library("aTSA") # for checking stationary test
```

```
##
## Attaching package: 'aTSA'
```

```
## The following object is masked from 'package:graphics':
##
##     identify
```

```
library("lattice") # for graphing
library("ggplot2") # for graphing
```

Transforming raw data into time series

```
TSData1 <- ts(TData1, frequency=12, start=c(1880,1))
```

Viewing the data and its properties.

```
TSData1
```

```
##      Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
## 1880 -29 -19 -17 -27 -13 -28 -22  -6 -16 -15 -18 -20
## 1881  -8 -13   2  -2  -3 -27  -5  -1  -8 -18 -25 -14
## 1882  10  10   2 -19 -17 -24  -9   5   0 -21 -20 -24
## 1883 -32 -41 -17 -23 -24 -11  -7 -12 -18 -11 -19 -17
## 1884 -17 -11 -33 -35 -31 -37 -33 -25 -22 -22 -30 -28
## 1885 -64 -29 -23 -44 -41 -50 -28 -27 -19 -19 -22  -5
## 1886 -41 -45 -41 -28 -27 -39 -15 -31 -18 -25 -26 -24
## 1887 -65 -47 -31 -37 -33 -20 -18 -27 -19 -32 -25 -37
## 1888 -42 -42 -47 -28 -21 -20  -8 -10  -7   2   0 -11
## 1889 -19  15   4   5  -3 -11  -5 -18 -18 -22 -30 -29
## 1890 -46 -48 -41 -38 -48 -27 -29 -35 -36 -23 -37 -29
## 1891 -45 -48 -14 -25 -17 -21 -21 -20 -13 -23 -36   1
## 1892 -24 -14 -35 -34 -24 -18 -26 -18 -24 -15 -49 -27
## 1893 -67 -50 -23 -32 -34 -22 -12 -22 -17 -15 -16 -36
## 1894 -54 -31 -21 -40 -29 -43 -31 -26 -22 -16 -24 -21
## 1895 -43 -42 -29 -22 -23 -24 -16 -15  -1 -10 -15 -11
## 1896 -22 -15 -29 -32 -19 -13  -5  -8  -3   6 -15 -11
## 1897 -22 -19 -13  -1   0 -13  -4  -3  -4  -9 -18 -26
## 1898  -7 -33 -56 -34 -36 -21 -23 -23 -20 -33 -35 -21
## 1899 -17 -39 -35 -21 -20 -26 -13  -4  -2   0  12 -26
## 1900 -39  -6   2 -14  -4 -13  -6  -3   2  10 -12 -13
## 1901 -28  -4   6  -5 -17  -9  -7 -11 -16 -27 -15 -28
## 1902 -18  -2 -28 -26 -29 -33 -25 -27 -20 -26 -35 -45
## 1903 -27  -5 -22 -39 -41 -43 -29 -43 -42 -41 -38 -46
## 1904 -63 -54 -45 -50 -50 -49 -47 -42 -46 -34 -16 -28
## 1905 -37 -58 -24 -36 -33 -30 -24 -20 -14 -22  -7 -20
## 1906 -29 -32 -14  -2 -20 -20 -25 -18 -24 -20 -38 -16
## 1907 -43 -50 -24 -40 -45 -42 -34 -36 -31 -23 -50 -49
## 1908 -44 -35 -57 -45 -39 -37 -33 -44 -32 -42 -50 -49
## 1909 -69 -46 -51 -59 -54 -52 -42 -30 -37 -39 -32 -53
## 1910 -44 -43 -47 -38 -33 -36 -31 -33 -36 -38 -55 -68
## 1911 -63 -60 -63 -55 -52 -47 -40 -42 -38 -25 -19 -24
## 1912 -27 -14 -37 -21 -20 -26 -40 -51 -47 -55 -37 -40
## 1913 -40 -43 -43 -36 -45 -45 -33 -31 -32 -33 -17  -3
## 1914   2 -13 -22 -27 -19 -21 -23 -13 -13  -5 -20  -9
## 1915 -19  -1  -9   8  -1 -14  -1 -14 -11 -21 -12 -24
## 1916 -19 -21 -29 -24 -26 -41 -31 -25 -28 -27 -42 -77
## 1917 -46 -55 -48 -38 -48 -39 -21 -25 -17 -34 -28 -70
## 1918 -43 -31 -20 -38 -36 -27 -19 -24 -13  -2 -14 -27
## 1919 -17 -20 -25 -18 -18 -26 -19 -18 -16 -14 -29 -33
```

```
## 1920 -14 -22  -6 -25 -24 -31 -30 -27 -18 -27 -32 -44
## 1921  -2 -20 -26 -34 -34 -29 -14 -22 -16  -5 -15 -17
## 1922 -31 -40 -12 -20 -33 -31 -24 -29 -27 -30 -15 -15
## 1923 -25 -35 -29 -37 -31 -22 -27 -28 -26 -10   5  -3
## 1924 -21 -25 -11 -33 -17 -26 -25 -32 -29 -34 -21 -40
## 1925 -32 -32 -21 -24 -29 -32 -27 -17 -12 -16   5  11
## 1926  22   8  13 -13 -23 -24 -19  -9 -10 -10  -6 -29
## 1927 -26 -19 -37 -31 -24 -26 -13 -18  -5   1  -3 -35
## 1928  -2 -10 -27 -28 -28 -40 -19 -24 -19 -18  -9 -19
## 1929 -46 -57 -34 -41 -38 -42 -32 -28 -23 -14 -14 -53
## 1930 -28 -23  -9 -24 -24 -17 -14  -9 -10  -8  15  -7
## 1931 -10 -20  -6 -20 -21  -5   2   0  -6   0 -10  -9
## 1932  14 -17 -19  -8 -22 -29 -23 -23 -12  -9 -26 -22
## 1933 -31 -30 -27 -23 -25 -31 -19 -22 -26 -23 -31 -45
## 1934 -25  -3 -28 -26 -10 -13 -10  -9 -15 -10   0  -7
## 1935 -37  12 -13 -34 -25 -22 -18 -16 -17  -7 -27 -20
## 1936 -28 -39 -24 -19 -15 -18  -5 -11  -5  -2  -4  -1
## 1937  -7   7 -16 -16  -6  -7  -3   4  15  10  11  -9
## 1938   3  -2   6   5  -7 -17  -8  -3   4  13   3 -23
## 1939 -12 -11 -19 -12  -7  -8  -5  -4   1  -3   7  41
## 1940 -13   6  12  16   6   6  11   1  12   8  13  20
## 1941  13  23   6  11  10   4  16  15   3  25  13  15
## 1942  29   7  12  14  15  12   3  -2   0   8  14  14
## 1943   1  23   3  15  11   1  15   4  12  31  27  29
## 1944  42  32  35  27  26  21  23  23  31  28  14   8
## 1945  15   5  11  25  11   3   8  26  22  22  10  -8
## 1946  17   6   0  12  -2 -15  -7  -9  -2  -6  -1 -27
## 1947  -8  -4   7   5  -6  -2  -3  -7 -14   7   0 -15
## 1948   6 -12 -23  -9   6  -5 -13  -9 -13  -6  -9 -20
## 1949  10 -16  -1  -7 -10 -25 -13  -9  -9  -3  -8 -15
## 1950 -27 -27  -8 -21 -12  -8  -9 -18 -10 -18 -34 -19
## 1951 -35 -43 -19 -11  -1  -7  -2   6   4   9   1  16
## 1952  16  13  -9   2  -5  -4   6   8   8  -3 -16  -2
## 1953   9  16  11  20   9   8   3   9   6   5  -5   4
## 1954 -27 -10 -12 -18 -20 -15 -16 -12  -7   0   9 -16
## 1955  12 -21 -35 -22 -20  -7  -7   7 -14  -5 -28 -32
## 1956 -16 -25 -22 -27 -27 -15 -11 -26 -19 -24 -16 -10
## 1957 -13  -6  -6  -2   7  14   0  14   8   0   6  16
## 1958  37  23   9   2   7  -6   3  -6  -5   3   2   0
## 1959   6   9  20  17   5   3   5  -1  -5  -9  -9  -2
## 1960  -2  13 -36 -17  -9   0  -3   0   4   8 -12  18
## 1961   7  19  10  15  14  11  -4   0   5   0   2 -15
## 1962   7  15  11   5  -6   5  -2  -3  -2  -4   6  -1
## 1963  -3  18 -15  -5 -10   3   9  23  20  14  16   0
## 1964  -6 -11 -24 -30 -26  -9  -7 -22 -28 -30 -21 -30
## 1965  -9 -17 -12 -20 -14 -10 -10  -1 -12  -4  -6  -6
## 1966 -16  -1   4 -14  -9   3   8 -11  -2 -15  -2  -6
## 1967  -7 -19   4  -6  12  -8  -1   3  -4   7  -6  -2
## 1968 -22 -14  21  -5 -10  -5 -10 -11 -19  11  -5 -14
## 1969 -11 -14   1  18  19   4  -2   3  10   9  12  27
## 1970  10  23   8   9  -5  -1  -4 -10  11   4   1 -14
## 1971  -3 -21 -19 -11  -7 -18 -13  -3   0  -6  -5  -8
## 1972 -26 -17   3  -1  -1   4   2  18   4   7   1  18
## 1973  27  31  25  26  26  16   9   2   6  12   5  -6
```

```
## 1974 -14 -28  -5 -10   0  -5  -3  11 -14  -7  -7 -10
## 1975   8   7  14   6  17   0  -2 -21  -2  -9 -16 -17
## 1976  -1  -6 -21  -9 -23 -15 -10 -16  -9 -26  -6   8
## 1977  17  21  24  26  30  25  23  20   2   5  20   5
## 1978   8  12  21  14   7  -2   7 -19   6   1  18  11
## 1979  15  -7  18  12   6  14   2  14  26  26  29  48
## 1980  30  42  29  33  35  18  28  26  22  19  29  20
## 1981  55  41  49  32  23  31  35  33  18  14  22  40
## 1982   9  16  -2   9  15   4  13   9  16  14  14  43
## 1983  53  41  42  30  36  19  15  31  38  16  31  17
## 1984  31  18  29  10  34   6  15  15  19  15   4  -6
## 1985  22  -5  17  10  18  17   2  15  15  13  10  16
## 1986  30  38  29  26  26  14  13  12   2  14  11  16
## 1987  36  46  17  26  26  36  47  28  40  33  26  48
## 1988  57  42  49  44  44  42  35  44  41  38  11  33
## 1989  16  35  36  33  17  15  33  35  37  33  19  36
## 1990  40  41  76  54  46  38  43  30  29  42  45  41
## 1991  42  50  35  51  39  54  51  42  48  31  31  33
## 1992  45  42  47  23  32  24  13  10  -1  10   4  22
## 1993  37  39  36  27  27  23  27  14  11  23   7  19
## 1994  31   3  26  40  29  41  31  22  34  42  46  36
## 1995  51  78  45  47  29  44  49  49  34  49  44  30
## 1996  27  49  34  36  29  26  35  50  26  21  41  40
## 1997  33  37  51  37  38  54  37  42  55  65  65  59
## 1998  61  89  62  64  71  77  70  69  46  46  49  58
## 1999  48  67  33  34  33  37  41  35  43  42  41  48
## 2000  26  59  59  59  40  43  42  43  43  29  33  30
## 2001  45  47  58  52  59  54  60  50  56  52  71  55
## 2002  74  76  91  59  65  55  61  55  64  57  59  42
## 2003  72  55  57  55  63  48  55  66  67  75  55  74
## 2004  58  70  66  62  42  42  27  45  53  67  72  52
## 2005  72  59  69  70  64  66  66  63  77  80  75  68
## 2006  57  71  64  49  47  64  55  71  64  70  73  78
## 2007  97  70  70  75  69  59  62  61  65  61  57  49
## 2008  27  36  75  53  51  48  60  45  64  67  67  55
## 2009  62  54  53  61  64  67  73  68  69  65  78  65
## 2010  73  79  92  87  75  64  61  65  61  71  79  49
## 2011  50  51  64  66  53  59  74  73  56  67  56  53
## 2012  45  49  57  69  76  62  56  64  75  79  74  52
## 2013  67  57  65  54  61  65  59  66  77  70  81  67
## 2014  74  50  77  78  86  66  58  82  90  86  68  79
```

```
str(TSData1)
```

```
##  Time-Series [1:1620] from 1880 to 2015: -29 -19 -17 -27 -13 -28 -22 -6 -16 -15 ...
```

```
summary(TSData1)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -77.000 -22.000  -5.000   1.599  20.000  97.000
```

```
frequency(TSData1)
```

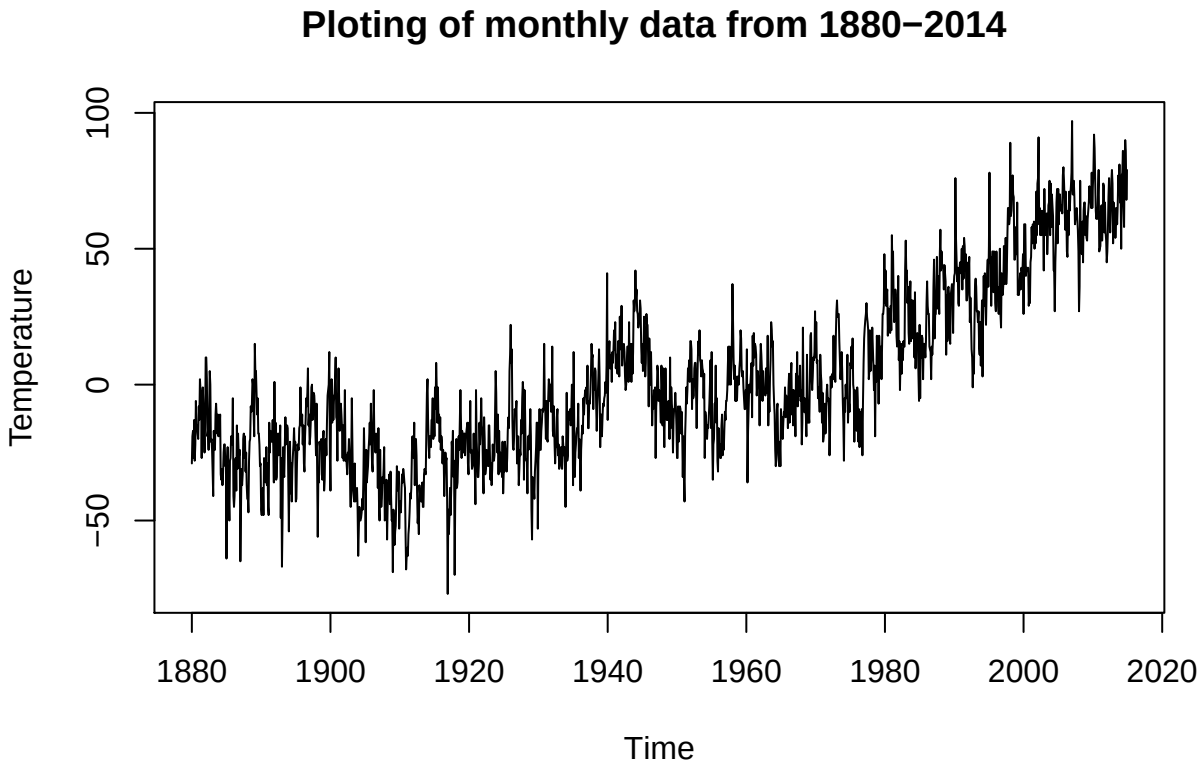
```
## [1] 12
```

```
aggregate(TSData1)
```

```
## Time Series:
## Start = 1880
## End = 2014
## Frequency = 1
##   [1] -230 -122 -107 -232 -324 -371 -360 -391 -234 -131 -437 -282 -308 -346
##  [15] -358 -251 -166 -132 -342 -191  -96 -161 -314 -416 -524 -325 -258 -467
##  [29] -507 -564 -502 -528 -415 -401 -183 -119 -390 -469 -294 -253 -300 -234
##  [43] -307 -268 -314 -226 -100 -236 -243 -422 -158 -105 -196 -333 -156 -224
##  [57] -171  -17  -26  -32   98  154  126  172  310  150  -34  -40 -107 -106
##  [71] -211  -82   14   95 -144 -172 -238   38   69   39  -36   64   31   70
##  [85] -244 -121  -61  -27  -83   76   32 -114   12  179  -92  -15 -134  218
##  [99]   84  203  331  393  160  369  190  150  231  409  480  345  525  507
## [113]  271  290  381  549  414  573  762  502  506  659  758  742  656  829
## [127]  763  795  648  779  856  722  758  789  894
```

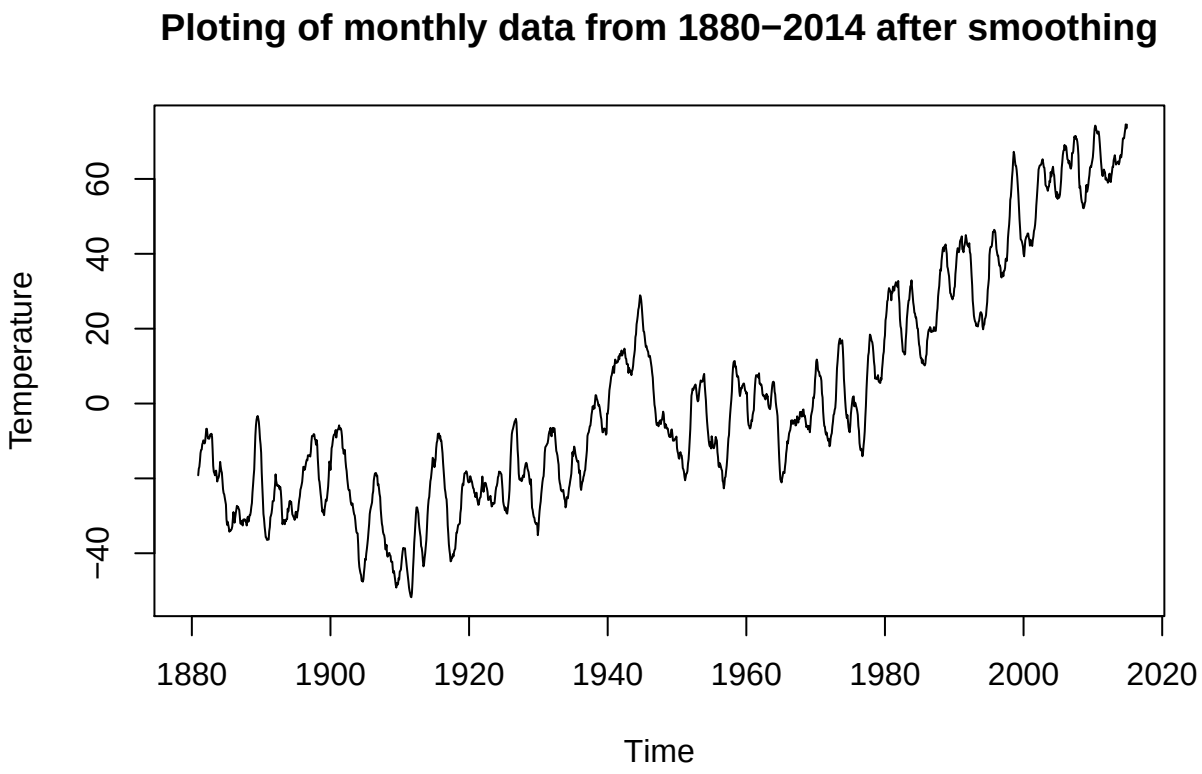
Plotting Time series data

```
plot.ts(TSData1, plot.type = c("multiple"), ylab="Temperature", main="Ploting of monthly data from 1880
```



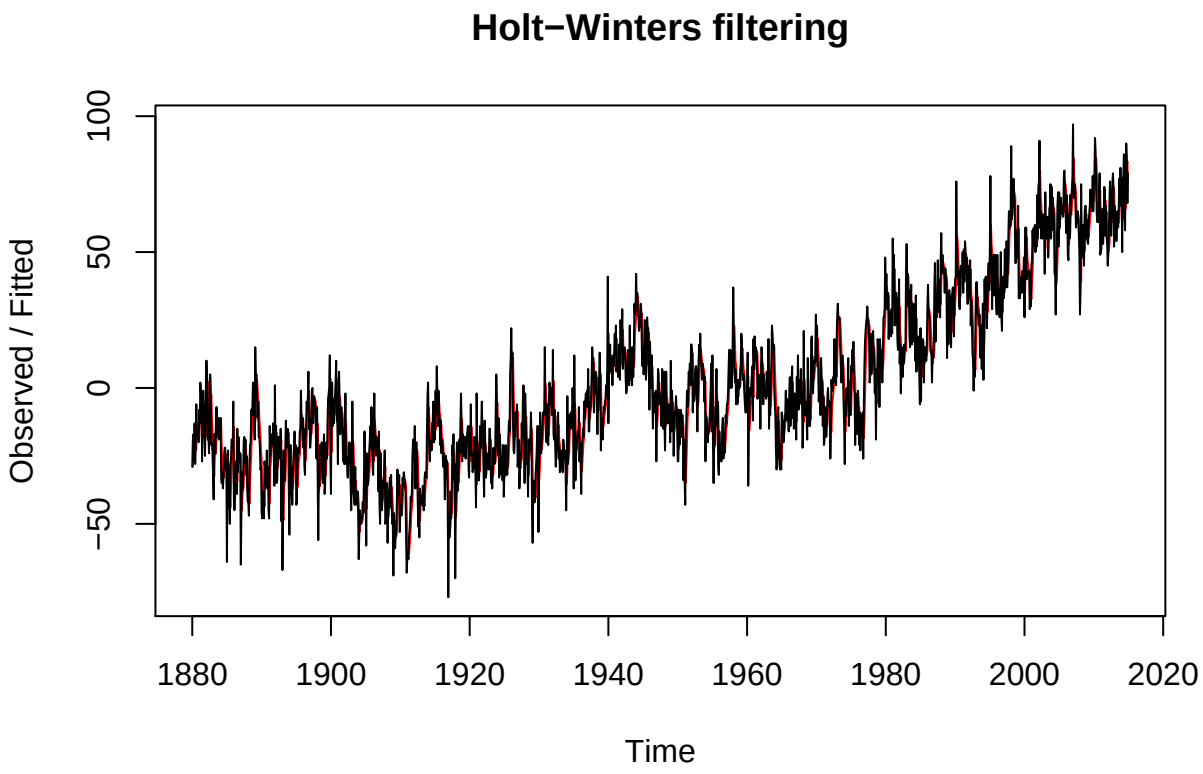
Further in order to see better pattern and trend in this time series we smooth time series data. For smoothing the order (span) of the simple moving average, using the parameter "n", needs to be specified. Since our data is monthly so order 12 is used in it.

```
plot.ts((SMA(TSData1,n=12)), ylab="Temperature", main="Ploting of monthly data from 1880-2014 after smo
```



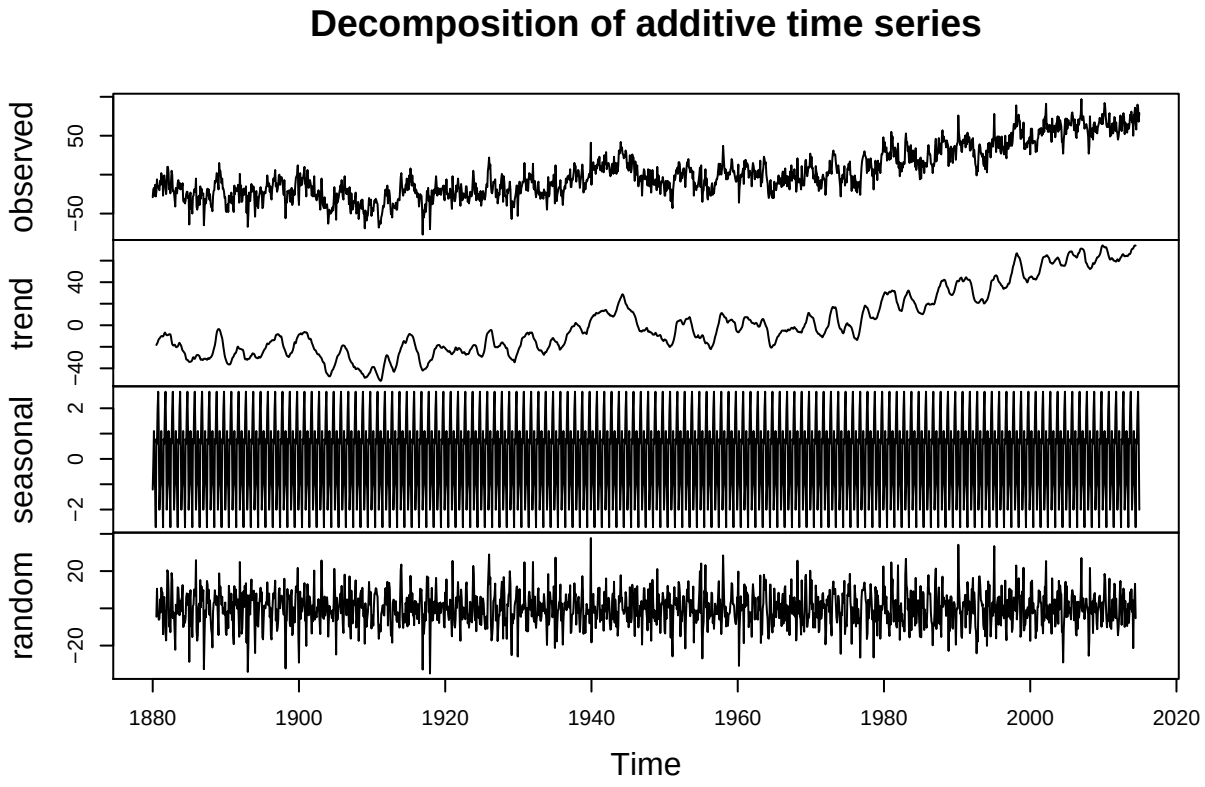
Smoothing factor can also be checked by using Holt winter function which plots data on observed vs fitted data. As shown below:

```
plot(HoltWinters(TSData1, beta = FALSE, gamma = FALSE))
```



Further to check the seasonality effect the data is decomposed as shown below:

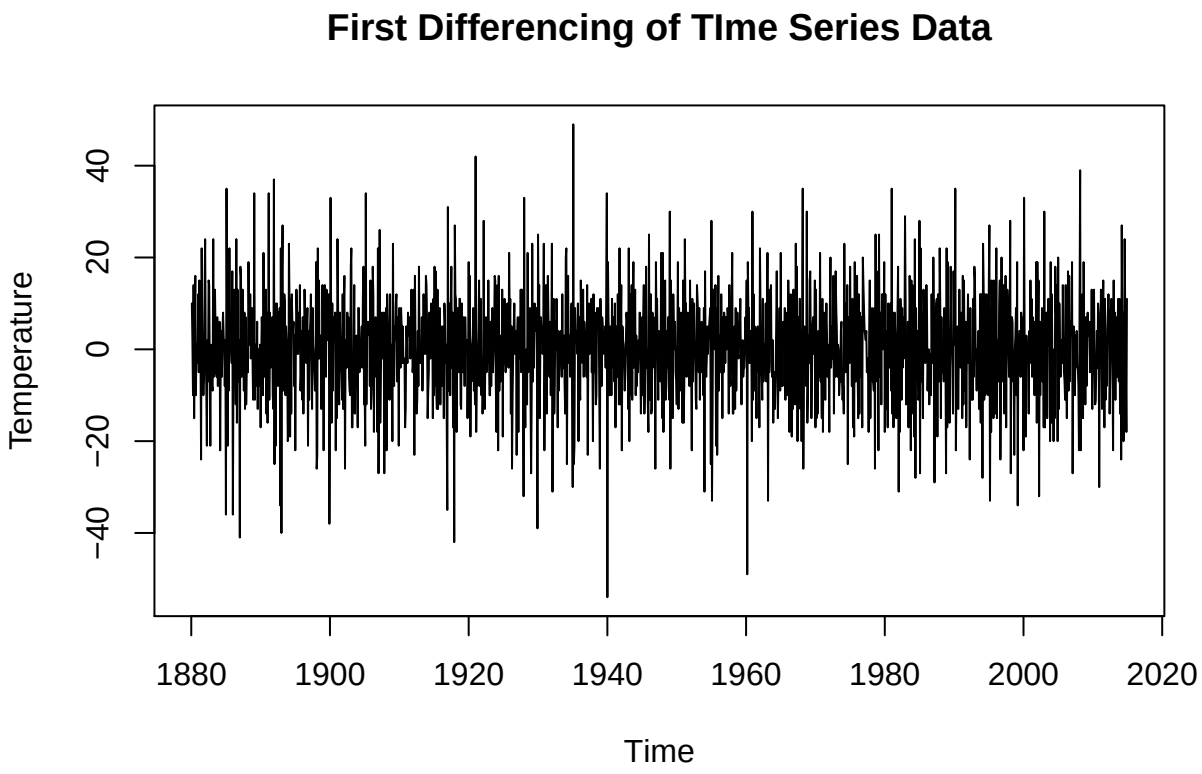
```
DTSData1<-decompose(TSData1)
plot(DTSData1)
```



After decomposing we see that there is a seasonality and an upward trend in all the observed, trend, and other two simple and smooth graphs. This clearly depicts that over the period of time global temperature is increasing.

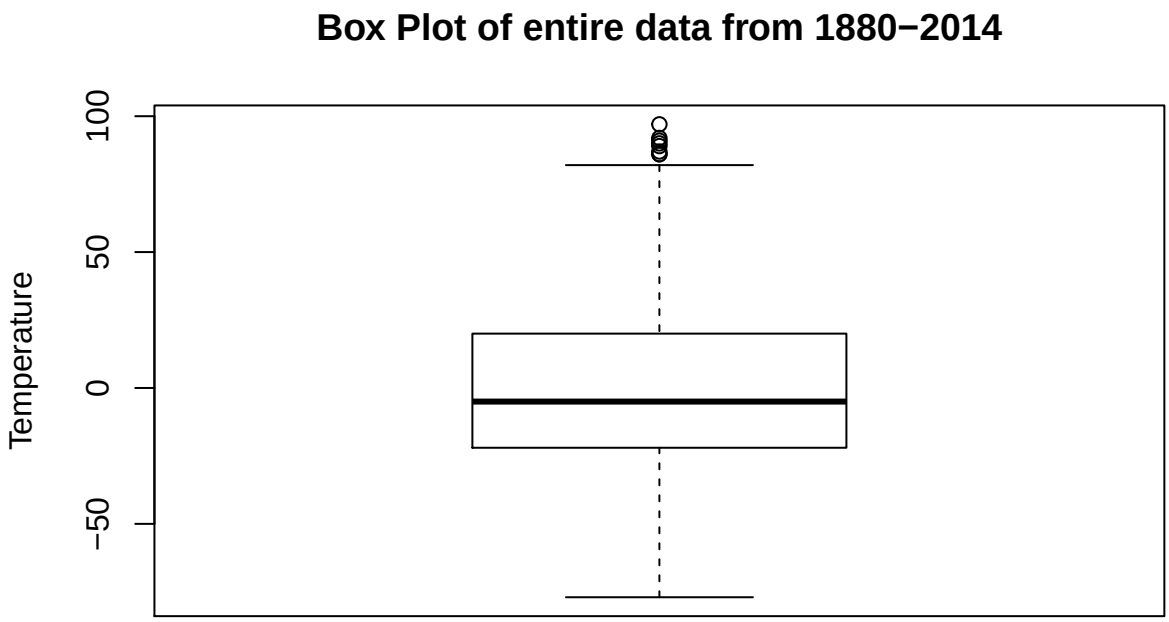
This seasonality and upward trend also depicts that data is non stationary. To correct for stationary we take first difference of data.

```
plot(diff(TSData1),ylab="Temperature", main="First Differencing of TIme Series Data")
```



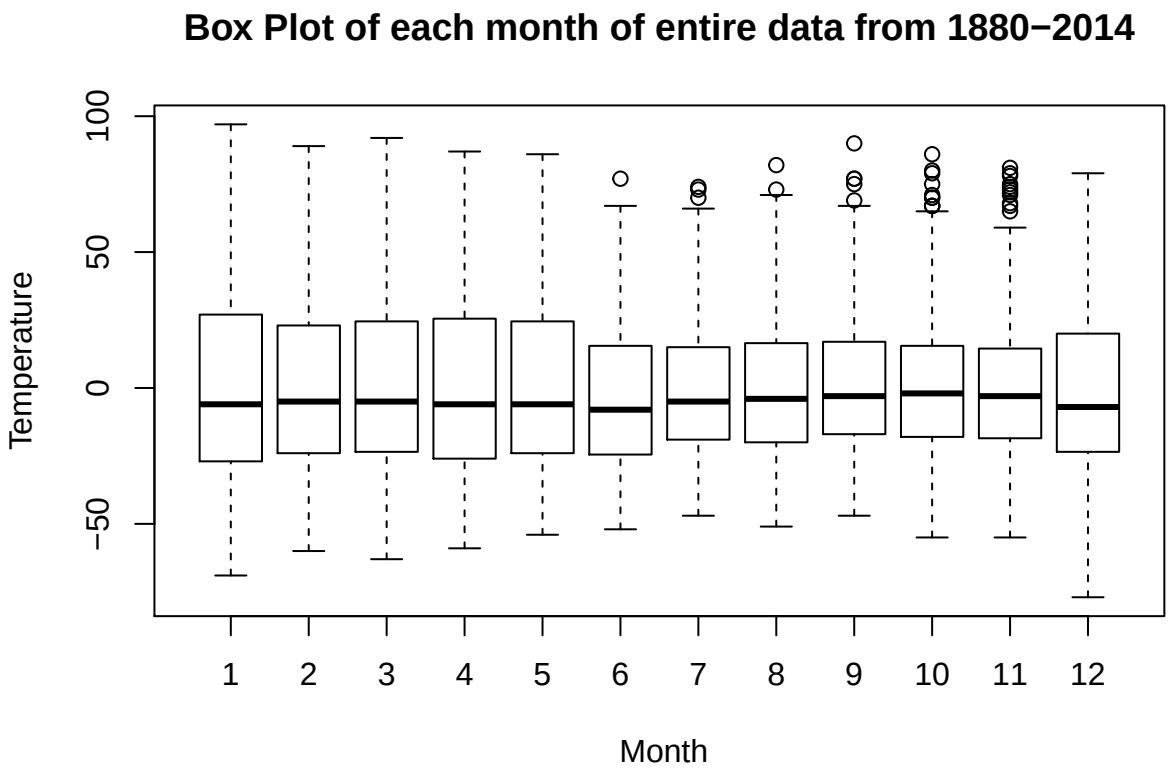
A further box plot is drawn for entire data check of any outliers and skewness in this data.

```
boxplot(TSData1, ylab="Temperature", main="Box Plot of entire data from 1880-2014")
```



Also, a box plot of each month for entire time series is also drawn to compare various month data.

```
boxplot(TSData1~cycle(TSData1), xlab=" Month", ylab="Temperature", main="Box Plot of each month of enti
```



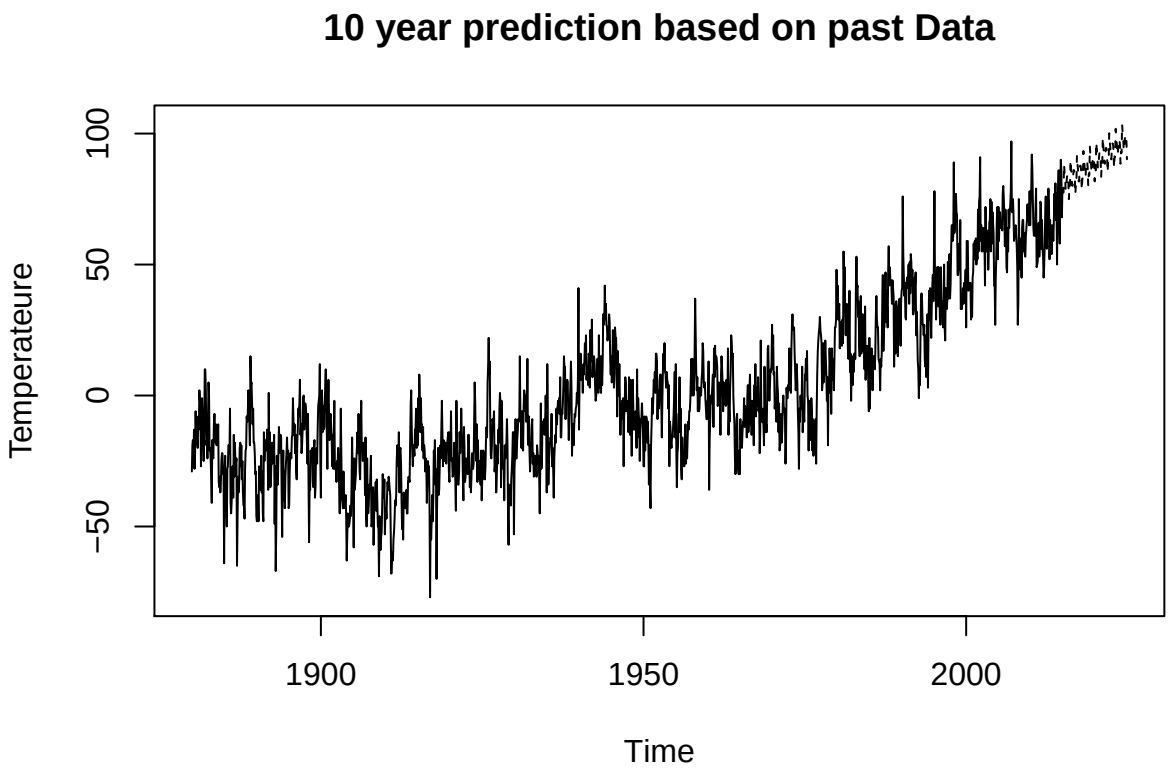
From all above analysis it is observed that there is an upward trend in the global temperature. Further if we want to predict what will be the temperature in coming 10 years based on our past data, it can be dome in many ways. Holt winter function is used here for making short term prediction. As shown below:

```
library("forecast")
```

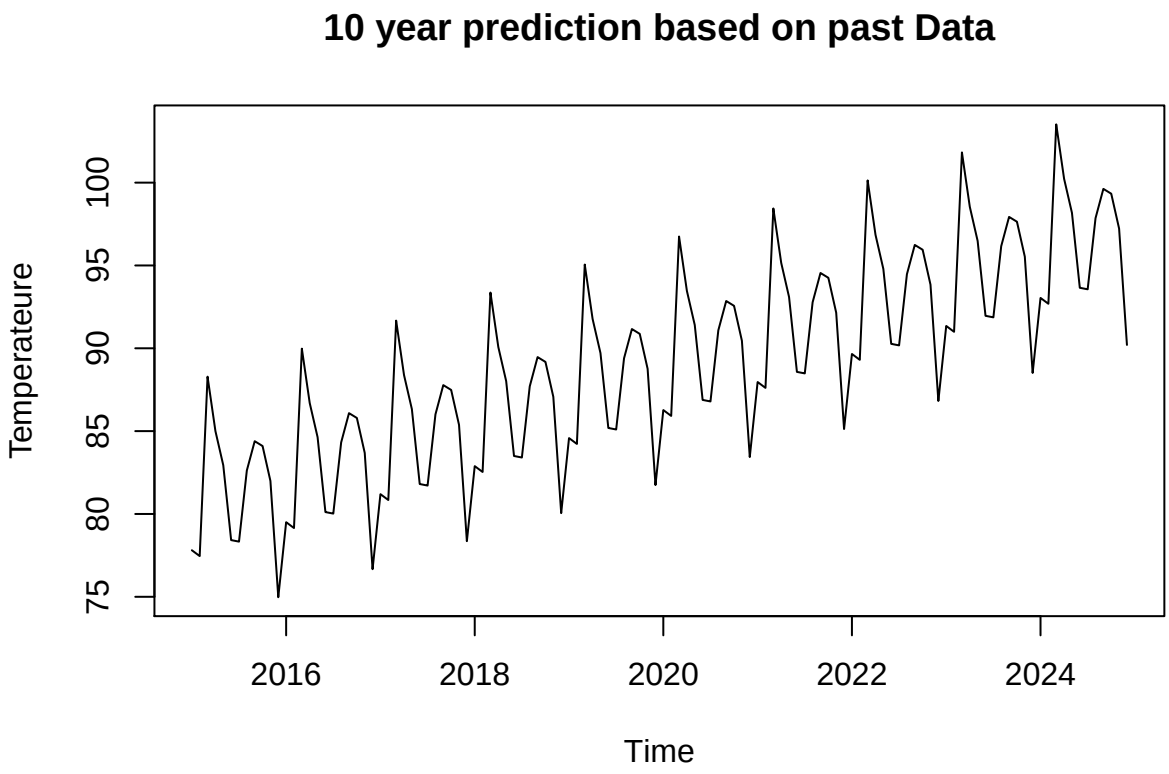
```
##
## Attaching package: 'forecast'
```

```
## The following object is masked from 'package:aTSA':
##
##     forecast
```

```
# Creating Holwinter object
TSData1.hw<- HoltWinters(TSData1)
# predicting 10 years ahead from this Data
TSData1.predict<-predict(TSData1.hw, n.ahead = 10*12)
# continuing past data and prediucted data
ts.plot(TSData1, TSData1.predict, lty=1:2, ylab="Temperateure",main="10 year prediction based on past Da
```




```
# ploting only predicted data
ts.plot(TSData1.predict, ylab="Temperateure",main="10 year prediction based on past Data")
```



Based on our prediction model we can also say that there is a constant rise in global temperature and if corrective measures are not taken then it is going to affect entire world.

Step 2 (Analysis of Quarterly Data)

For analysis of quarterly data we simply imported raw data provided in CSV format. As shwon below:

```
# Read File
RData1 <- read.csv(file = "E:/onedrive/University of Nebraska Omaha/Courses/Data Visualization/Assignmen
```

In order to check whether imported data is in the proper format or not str function is used to check the details:

```
str(RData1)
```

```
## 'data.frame':    136 obs. of  19 variables:
##  $ Year: int  1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 ...
##  $ Jan : int  -29 -8 10 -32 -17 -64 -41 -65 -42 -19 ...
##  $ Feb : int  -19 -13 10 -41 -11 -29 -45 -47 -42 15 ...
##  $ Mar : int  -17 2 2 -17 -33 -23 -41 -31 -47 4 ...
##  $ Apr : int  -27 -2 -19 -23 -35 -44 -28 -37 -28 5 ...
##  $ May : int  -13 -3 -17 -24 -31 -41 -27 -33 -21 -3 ...
##  $ Jun : int  -28 -27 -24 -11 -37 -50 -39 -20 -20 -11 ...
##  $ Jul : chr  "-22" "-5" "-9" "-7" ...
##  $ Aug : chr  "-6" "-1" "5" "-12" ...
##  $ Sep : chr  "-16" "-8" "0" "-18" ...
```

```
##  $ Oct : chr  "-15" "-18" "-21" "-11" ...
##  $ Nov : chr  "-18" "-25" "-20" "-19" ...
##  $ Dec : chr  "-20" "-14" "-24" "-17" ...
##  $ J.D : chr  "-19" "-10" "-9" "-19" ...
##  $ D.N : chr  "***" "-11" "-8" "-20" ...
##  $ DJF : chr  "****" "-13" "2" "-32" ...
##  $ MAM : int  -19 -1 -11 -22 -33 -36 -32 -33 -32 2 ...
##  $ JJA : chr  "-19" "-11" "-9" "-10" ...
##  $ SON : chr  "-16" "-17" "-14" "-16" ...
```

It is observed that some of the variables are in int format and some are in character format. For this data analysis, all the variables need to be converted into a numeric format. As shown below:

```
RData1$Jan<-as.numeric(RData1$Jan)
RData1$Feb<-as.numeric(RData1$Feb)
RData1$Mar<-as.numeric(RData1$Mar)
RData1$Apr<-as.numeric(RData1$Apr)
RData1$May<-as.numeric(RData1$May)
RData1$Jun<-as.numeric(RData1$Jun)
RData1$Jul<-as.numeric(RData1$Jul)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$Aug<-as.numeric(RData1$Aug)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$Sep<-as.numeric(RData1$Sep)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$Oct<-as.numeric(RData1$Oct)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$Nov<-as.numeric(RData1$Nov)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$Dec<-as.numeric(RData1$Dec)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$J.D<-as.numeric(RData1$J.D)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$D.N<-as.numeric(RData1$D.N)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$DJF<-as.numeric(RData1$DJF)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$MAM<-as.numeric(RData1$MAM)
RData1$JJA<-as.numeric(RData1$JJA)
```

```
## Warning: NAs introduced by coercion
```

```
RData1$SON<-as.numeric(RData1$SON)
```

```
## Warning: NAs introduced by coercion
```

Now viewing the data and its properties.

```
str(RData1)
```

```
## 'data.frame':    136 obs. of  19 variables:
##  $ Year: int  1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 ...
##  $ Jan : num  -29 -8 10 -32 -17 -64 -41 -65 -42 -19 ...
##  $ Feb : num  -19 -13 10 -41 -11 -29 -45 -47 -42 15 ...
##  $ Mar : num  -17 2 2 -17 -33 -23 -41 -31 -47 4 ...
##  $ Apr : num  -27 -2 -19 -23 -35 -44 -28 -37 -28 5 ...
##  $ May : num  -13 -3 -17 -24 -31 -41 -27 -33 -21 -3 ...
##  $ Jun : num  -28 -27 -24 -11 -37 -50 -39 -20 -20 -11 ...
##  $ Jul : num  -22 -5 -9 -7 -33 -28 -15 -18 -8 -5 ...
##  $ Aug : num  -6 -1 5 -12 -25 -27 -31 -27 -10 -18 ...
##  $ Sep : num  -16 -8 0 -18 -22 -19 -18 -19 -7 -18 ...
##  $ Oct : num  -15 -18 -21 -11 -22 -19 -25 -32 2 -22 ...
##  $ Nov : num  -18 -25 -20 -19 -30 -22 -26 -25 0 -30 ...
##  $ Dec : num  -20 -14 -24 -17 -28 -5 -24 -37 -11 -29 ...
##  $ J.D : num  -19 -10 -9 -19 -27 -31 -30 -33 -20 -11 ...
##  $ D.N : num  NA -11 -8 -20 -26 -33 -28 -32 -22 -9 ...
##  $ DJF : num  NA -13 2 -32 -15 -41 -30 -46 -40 -5 ...
##  $ MAM : num  -19 -1 -11 -22 -33 -36 -32 -33 -32 2 ...
##  $ JJA : num  -19 -11 -9 -10 -32 -35 -28 -21 -13 -11 ...
##  $ SON : num  -16 -17 -14 -16 -25 -20 -23 -26 -2 -23 ...
```

```
summary(RData1)
```

```
##       Year           Jan                Feb               Mar         
##  Min.   :1880   Min.   :-69.0000   Min.   :-60.000   Min.   :-63.000  
##  1st Qu.:1914   1st Qu.:-27.0000   1st Qu.:-23.500   1st Qu.:-23.250  
##  Median :1948   Median : -4.5000   Median : -5.000   Median : -3.500  
##  Mean   :1948   Mean   :  0.6103   Mean   :  1.846   Mean   :  3.147  
##  3rd Qu.:1981   3rd Qu.: 27.0000   3rd Qu.: 25.000   3rd Qu.: 25.250  
##  Max.   :2015   Max.   : 97.0000   Max.   : 89.000   Max.   : 92.000  
##                                                                       
##       Apr               May                Jun               Jul         
##  Min.   :-59.000   Min.   :-54.0000   Min.   :-52.0000   Min.   :-47.000  
##  1st Qu.:-26.000   1st Qu.:-24.0000   1st Qu.:-24.2500   1st Qu.:-19.000  
##  Median : -5.500   Median : -6.0000   Median : -7.5000   Median : -5.000  
##  Mean   :  1.162   Mean   :  0.9044   Mean   : -0.5735   Mean   :  2.296  
##  3rd Qu.: 26.000   3rd Qu.: 26.0000   3rd Qu.: 16.2500   3rd Qu.: 15.000  
##  Max.   : 87.000   Max.   : 86.0000   Max.   : 80.0000   Max.   : 74.000  
##                                                          NA's   :1        
##       Aug               Sep                Oct               Nov          
##  Min.   :-51.000   Min.   :-47.000   Min.   :-55.000   Min.   :-55.00  
##  1st Qu.:-20.000   1st Qu.:-17.000   1st Qu.:-18.000   1st Qu.:-18.50  
##  Median : -4.000   Median : -3.000   Median : -2.000   Median : -3.00  
##  Mean   :  2.341   Mean   :  3.778   Mean   :  4.519   Mean   :  2.77  
##  3rd Qu.: 16.500   3rd Qu.: 17.000   3rd Qu.: 15.500   3rd Qu.: 14.50  
##  Max.   : 82.000   Max.   : 90.000   Max.   : 86.000   Max.   : 81.00  
##  NA's   :1         NA's   :1         NA's   :1         NA's   :1       
##       Dec                J.D               D.N               DJF          
##  Min.   :-77.00000   Min.   :-47.00   Min.   :-48.000   Min.   :-63.0000  
##  1st Qu.:-23.50000   1st Qu.:-20.00   1st Qu.:-22.000   1st Qu.:-24.0000  
##  Median : -7.00000   Median : -8.00   Median : -7.500   Median : -8.0000  
```

```
##  Mean   : -0.03704   Mean   :  1.63   Mean   :  1.687   Mean   :  0.9259
##  3rd Qu.: 20.00000   3rd Qu.: 17.50   3rd Qu.: 18.750   3rd Qu.: 25.0000
##  Max.   : 79.00000   Max.   : 75.00   Max.   : 74.000   Max.   : 83.0000
##  NA's   :1           NA's   :1        NA's   :2         NA's   :1
##       MAM              JJA              SON
##  Min.   :-56.000   Min.   :-46.000   Min.   :-47.000
##  1st Qu.:-25.000   1st Qu.:-21.000   1st Qu.:-17.000
##  Median : -5.000   Median : -7.000   Median : -3.000
##  Mean   :  1.728   Mean   :  1.185   Mean   :  3.733
##  3rd Qu.: 25.500   3rd Qu.: 13.000   3rd Qu.: 15.500
##  Max.   : 85.000   Max.   : 72.000   Max.   : 81.000
##                    NA's   :1         NA's   :1
```

```
head(RData1)
```

```
##   Year Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec J.D D.N DJF MAM JJA
## 1 1880 -29 -19 -17 -27 -13 -28 -22  -6 -16 -15 -18 -20 -19  NA  NA -19 -19
## 2 1881  -8 -13   2  -2  -3 -27  -5  -1  -8 -18 -25 -14 -10 -11 -13  -1 -11
## 3 1882  10  10   2 -19 -17 -24  -9   5   0 -21 -20 -24  -9  -8   2 -11  -9
## 4 1883 -32 -41 -17 -23 -24 -11  -7 -12 -18 -11 -19 -17 -19 -20 -32 -22 -10
## 5 1884 -17 -11 -33 -35 -31 -37 -33 -25 -22 -22 -30 -28 -27 -26 -15 -33 -32
## 6 1885 -64 -29 -23 -44 -41 -50 -28 -27 -19 -19 -22  -5 -31 -33 -41 -36 -35
##   SON
## 1 -16
## 2 -17
## 3 -14
## 4 -16
## 5 -25
## 6 -20
```

```
tail(RData1)
```

```
##     Year Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec J.D D.N DJF MAM
## 131 2010  73  79  92  87  75  64  61  65  61  71  79  49  71  73  72  85
## 132 2011  50  51  64  66  53  59  74  73  56  67  56  53  60  60  50  61
## 133 2012  45  49  57  69  76  62  56  64  75  79  74  52  63  63  49  67
## 134 2013  67  57  65  54  61  65  59  66  77  70  81  67  66  64  58  60
## 135 2014  74  50  77  78  86  66  58  82  90  86  68  79  75  74  64  81
## 136 2015  82  88  90  74  76  80  NA  NA  NA  NA  NA  NA  NA  NA  83  80
##     JJA SON
## 131  63  71
## 132  69  60
## 133  61  76
## 134  63  76
## 135  69  81
## 136  NA  NA
```

```
sum(is.na(RData1))
```

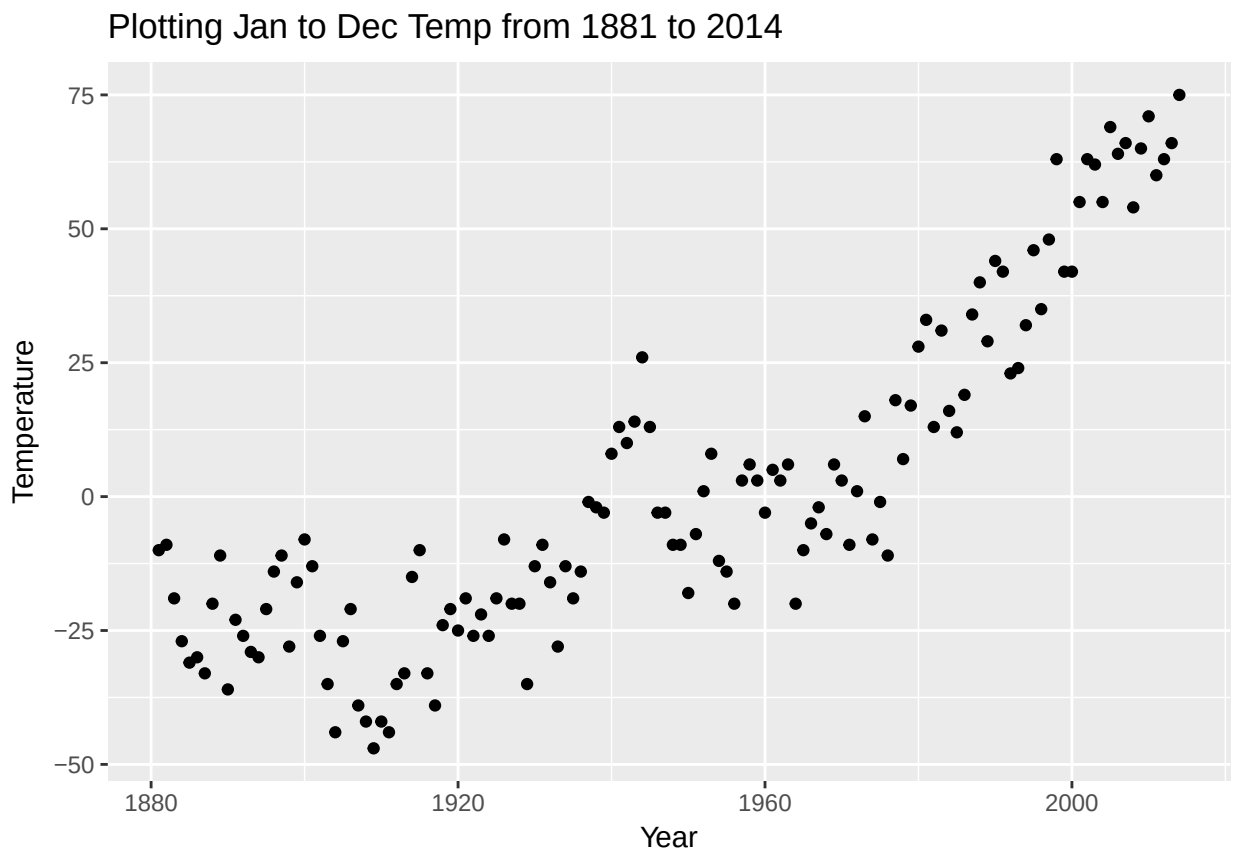
```
## [1] 12
```

Since year 1880 and 2015 have some missing values so these two years are droped from the dataset.

```
RData1<-RData1[-c(1,136),]
```

Ploting temperature from Jan to Dec

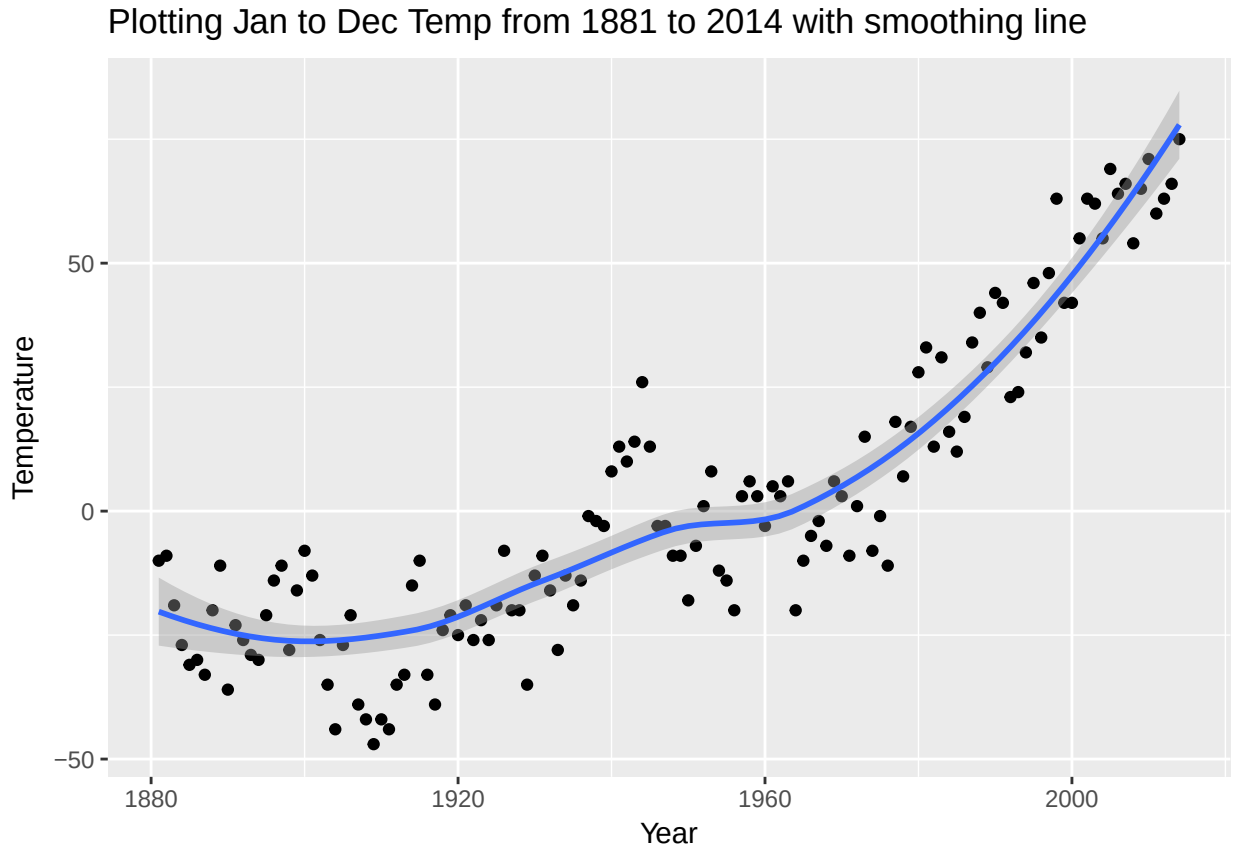
```
ggplot(RData1, aes(Year, J.D))+
  geom_point()+
  ggtitle("Plotting Jan to Dec Temp from 1881 to 2014")+
  labs(y="Temperature")
```



Ploting the same Jan to Dec Temp with smooth line.

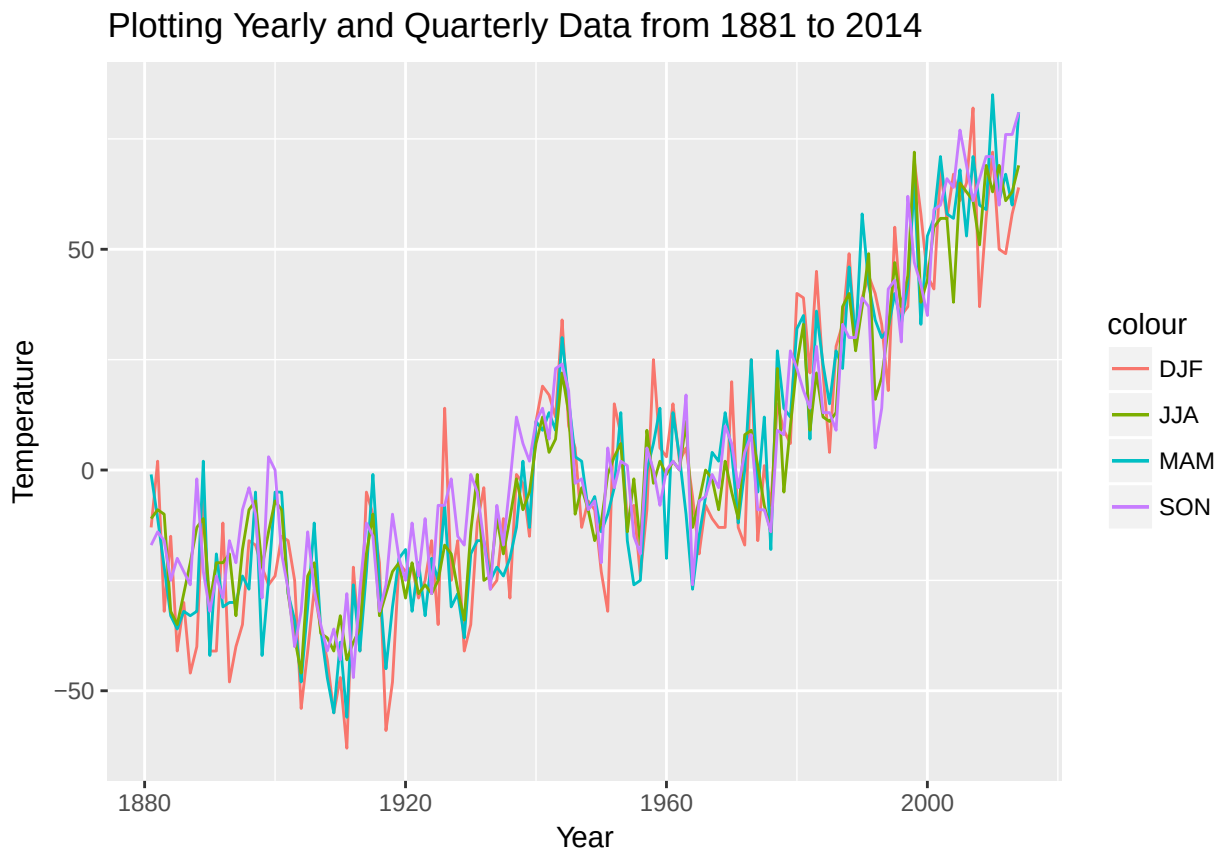
```
ggplot(RData1, aes(Year, J.D))+
  geom_point()+
  geom_smooth()+
  ggtitle("Plotting Jan to Dec Temp from 1881 to 2014 with smoothing line")+
  labs(y="Temperature")
```

```
## `geom_smooth()` using method = 'loess'
```



Plotting Quaterly Data

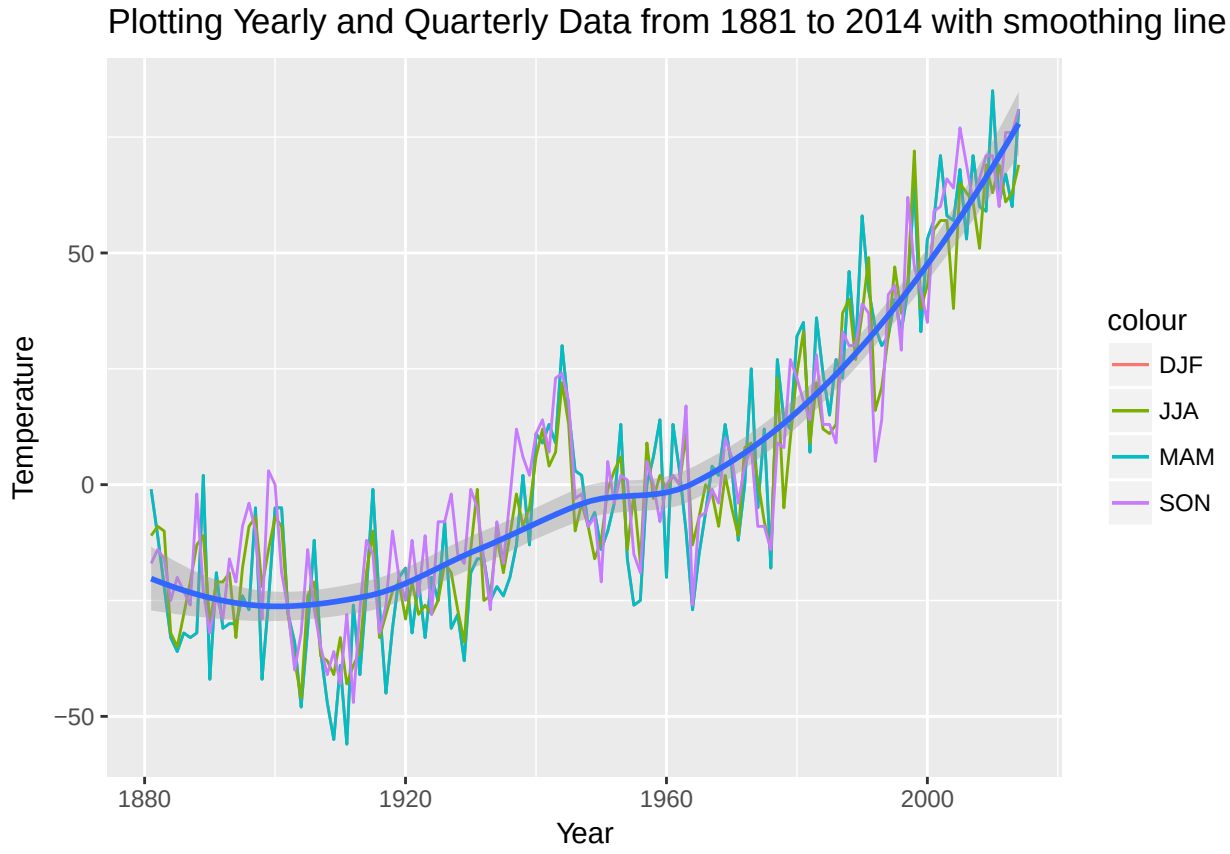
```
ggplot(RData1, aes(Year))+
  geom_line(aes(y=DJF, colour="DJF"))+
  geom_line(aes(y=MAM, colour="MAM"))+
  geom_line(aes(y=JJA, colour="JJA"))+
  geom_line(aes(y=SON, colour="SON"))+
  ggtitle("Plotting Yearly and Quarterly Data from 1881 to 2014")+
  labs(y="Temperature")
```



Plotting Yearly and Quarterly Data with smoothing line

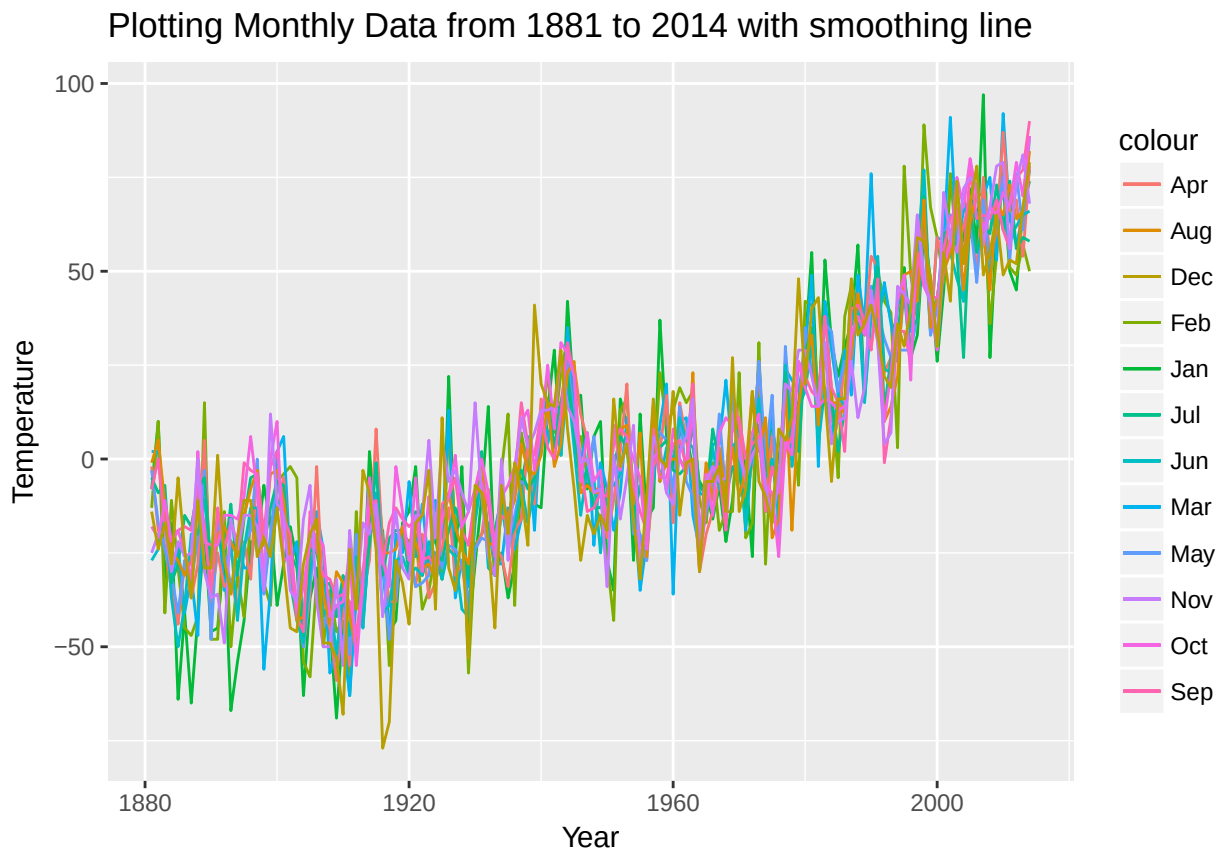
```
ggplot(RData1, aes(Year, J.D))+
  geom_line(aes(y=MAM, colour="DJF"))+
  geom_line(aes(y=MAM, colour="MAM"))+
  geom_line(aes(y=JJA, colour="JJA"))+
  geom_line(aes(y=SON, colour="SON"))+
  geom_smooth()+
  ggtitle("Plotting Yearly and Quarterly Data from 1881 to 2014 with smoothing line")+
  labs(y="Temperature")
```

```
## `geom_smooth()` using method = 'loess'
```



Plotting Monthly Data

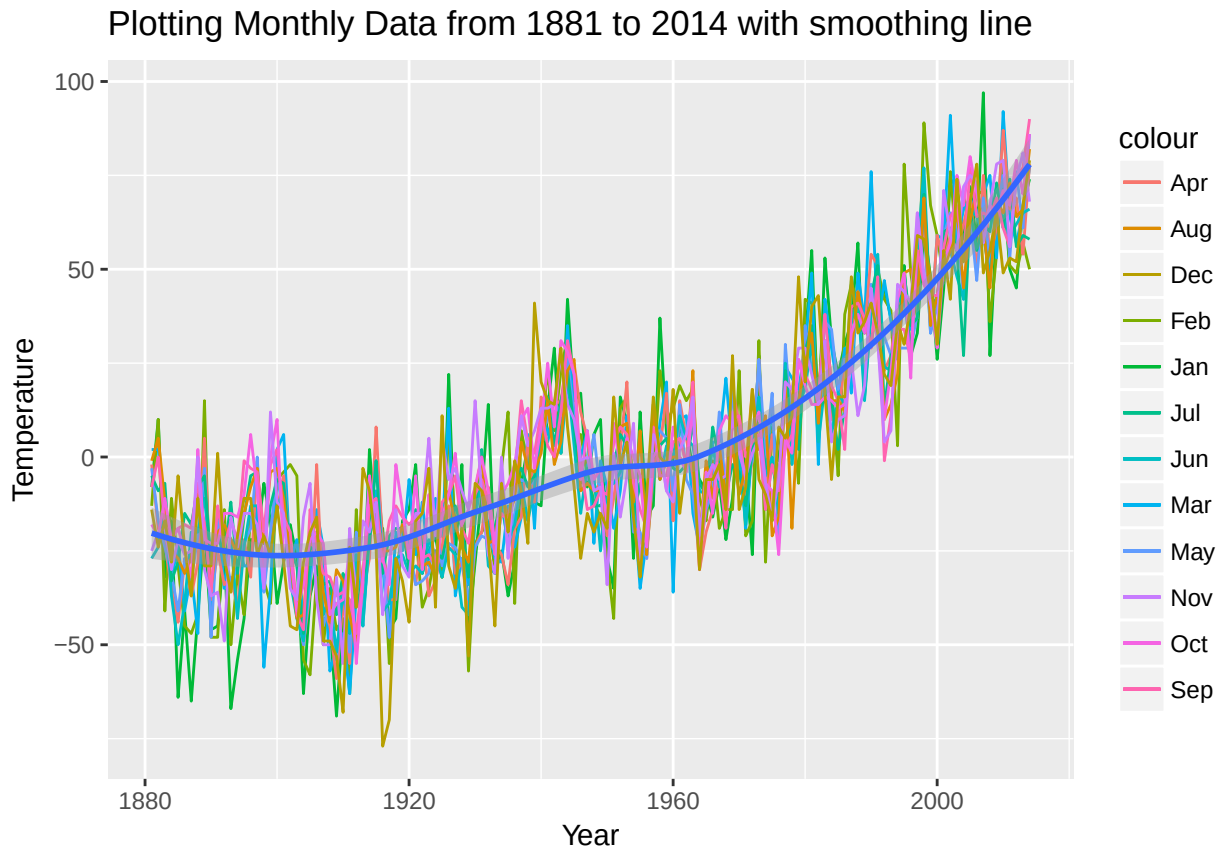
```
ggplot(RData1, aes(Year))+
  geom_line(aes(y=Jan, colour="Jan"))+
  geom_line(aes(y=Feb, colour="Feb"))+
  geom_line(aes(y=Mar, colour="Mar"))+
  geom_line(aes(y=Apr, colour="Apr"))+
  geom_line(aes(y=May, colour="May"))+
  geom_line(aes(y=Jun, colour="Jun"))+
  geom_line(aes(y=Jul, colour="Jul"))+
  geom_line(aes(y=Aug, colour="Aug"))+
  geom_line(aes(y=Sep, colour="Sep"))+
  geom_line(aes(y=Oct, colour="Oct"))+
  geom_line(aes(y=Nov, colour="Nov"))+
  geom_line(aes(y=Dec, colour="Dec"))+
  ggtitle("Plotting Monthly Data from 1881 to 2014 with smoothing line")+
  labs(y="Temperature")
```



Plotting Monthly Data with smoothing line

```
ggplot(RData1, aes(Year, J.D))+
  geom_line(aes(y=Jan, colour="Jan"))+
  geom_line(aes(y=Feb, colour="Feb"))+
  geom_line(aes(y=Mar, colour="Mar"))+
  geom_line(aes(y=Apr, colour="Apr"))+
  geom_line(aes(y=May, colour="May"))+
  geom_line(aes(y=Jun, colour="Jun"))+
  geom_line(aes(y=Jul, colour="Jul"))+
  geom_line(aes(y=Aug, colour="Aug"))+
  geom_line(aes(y=Sep, colour="Sep"))+
  geom_line(aes(y=Oct, colour="Oct"))+
  geom_line(aes(y=Nov, colour="Nov"))+
  geom_line(aes(y=Dec, colour="Dec"))+
  geom_smooth()+
  ggtitle("Plotting Monthly Data from 1881 to 2014 with smoothing line")+
  labs(y="Temperature")
```

```
## `geom_smooth()` using method = 'loess'
```



From All of the above graphs, it is observed that there is an upward trend in yearly, quarterly and monthly data. Hence step 2 also supports that there is an upward trend in global temperature.

Step 3 (Analysis of Hemispheres Data)

In this step temperature data of different global horizons from north pole to south pole is used to analyze and check the upward trend in global temperature.

```
# Read File
RData2 <- read.csv(file = "E:/onedrive/University of Nebraska Omaha/Courses/Data Visualization/Assignmen
```

In order to check whether imported data is in the proper format or not str function is used to check the details:

```
str(RData2)
```

```
## 'data.frame':    135 obs. of  15 variables:
##  $ Year    : int  1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 ...
##  $ Glob    : int  -19 -10 -9 -19 -27 -31 -30 -33 -20 -11 ...
##  $ NHem    : int  -33 -18 -17 -30 -42 -41 -39 -37 -22 -16 ...
##  $ SHem    : int  -5 -2 -1 -8 -12 -21 -21 -28 -17 -6 ...
##  $ X24N.90N: int  -38 -27 -21 -34 -56 -61 -49 -46 -42 -25 ...
##  $ X24S.24N: int  -16 -2 -10 -22 -17 -17 -24 -27 7 4 ...
##  $ X90S.24S: int  -5 -5 4 -2 -11 -20 -20 -26 -33 -17 ...
##  $ X64N.90N: int  -89 -54 -125 -28 -127 -119 -124 -158 -141 -82 ...
##  $ X44N.64N: int  -54 -40 -20 -57 -58 -70 -43 -52 -43 -13 ...
```

```
##  $ X24N.44N: int  -22 -14 -3 -20 -41 -43 -38 -21 -22 -21 ...
##  $ EQU.24N : int  -26 -5 -12 -25 -21 -11 -24 -24 7 -3 ...
##  $ X24S.EQU: int  -5 2 -8 -19 -14 -23 -24 -31 8 11 ...
##  $ X44S.24S: int  -2 -6 3 -1 -15 -27 -18 -24 -30 -16 ...
##  $ X64S.44S: int  -8 -3 8 0 -5 -7 -21 -29 -38 -17 ...
##  $ X90S.64S: int  39 37 42 37 40 38 28 21 16 19 ...
```

Since all the imported data is in the proper format so no need to make any changes to the data structure.

Further to check if there are any missing in the data or not we used:

```
sum(is.na(RData2))
```

```
## [1] 0
```

Further, there are no missing values in data and this data is from 1880 to 2014 so entire data is used in the evaluation.

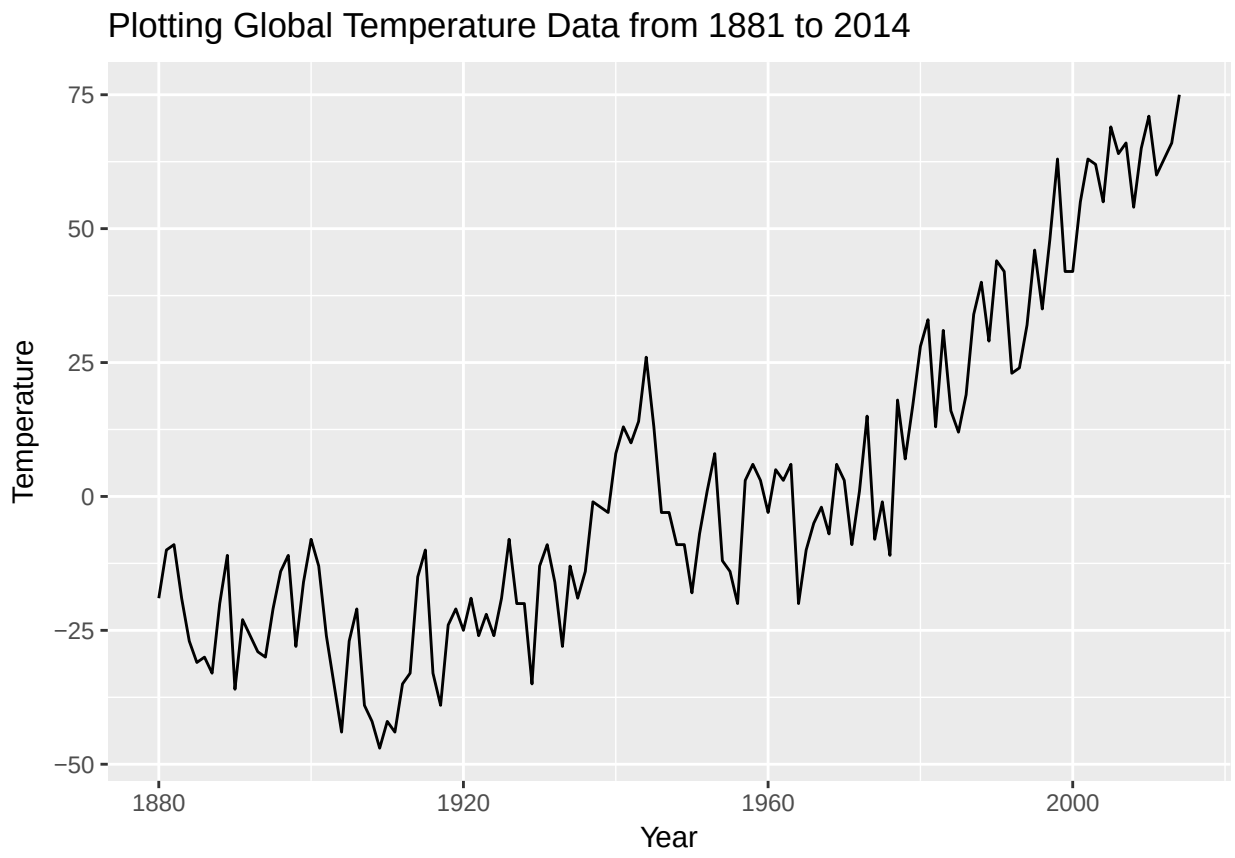
Data Summary

```
summary(RData2)
```

```
##       Year           Glob             NHem              SHem
##  Min.   :1880   Min.   :-47.00   Min.   :-52.000   Min.   :-47.00000
##  1st Qu.:1914   1st Qu.:-20.00   1st Qu.:-21.500   1st Qu.:-22.50000
##  Median :1947   Median : -8.00   Median : -2.000   Median : -9.00000
##  Mean   :1947   Mean   :  1.63   Mean   :  3.326   Mean   : -0.07407
##  3rd Qu.:1980   3rd Qu.: 17.50   3rd Qu.: 16.000   3rd Qu.: 25.00000
##  Max.   :2014   Max.   : 75.00   Max.   : 91.000   Max.   : 59.00000
##     X24N.90N          X24S.24N          X90S.24S          X64N.90N
##  Min.   :-61.000   Min.   :-61.000   Min.   :-48.000   Min.   :-158.000
##  1st Qu.:-26.000   1st Qu.:-22.000   1st Qu.:-26.000   1st Qu.: -47.500
##  Median :  2.000   Median : -3.000   Median :-11.000   Median :   3.000
##  Mean   :  5.415   Mean   :  1.926   Mean   : -2.704   Mean   :   9.022
##  3rd Qu.: 21.000   3rd Qu.: 23.000   3rd Qu.: 21.000   3rd Qu.:  58.000
##  Max.   :110.000   Max.   : 72.000   Max.   : 58.000   Max.   : 211.000
##     X44N.64N          X24N.44N          EQU.24N
##  Min.   :-70.000   Min.   :-57.0000   Min.   :-70.00000
##  1st Qu.:-27.000   1st Qu.:-19.5000   1st Qu.:-23.50000
##  Median :  0.000   Median : -8.0000   Median : -3.00000
##  Mean   :  9.163   Mean   :  0.7111   Mean   :  0.08148
##  3rd Qu.: 34.500   3rd Qu.: 13.5000   3rd Qu.: 19.50000
##  Max.   :129.000   Max.   : 77.0000   Max.   : 72.00000
##     X24S.EQU          X44S.24S          X64S.44S          X90S.64S
##  Min.   :-55.000   Min.   :-43.0000   Min.   :-62.000   Min.   :-237.000
##  1st Qu.:-22.000   1st Qu.:-23.0000   1st Qu.:-27.500   1st Qu.: -41.000
##  Median : -3.000   Median : -9.0000   Median : -9.000   Median :   5.000
##  Mean   :  3.748   Mean   :  0.7926   Mean   : -7.593   Mean   :  -5.119
##  3rd Qu.: 29.500   3rd Qu.: 22.0000   3rd Qu.: 16.000   3rd Qu.:  37.500
##  Max.   : 81.000   Max.   : 76.0000   Max.   : 38.000   Max.   : 136.000
```

Ploting of global temperature

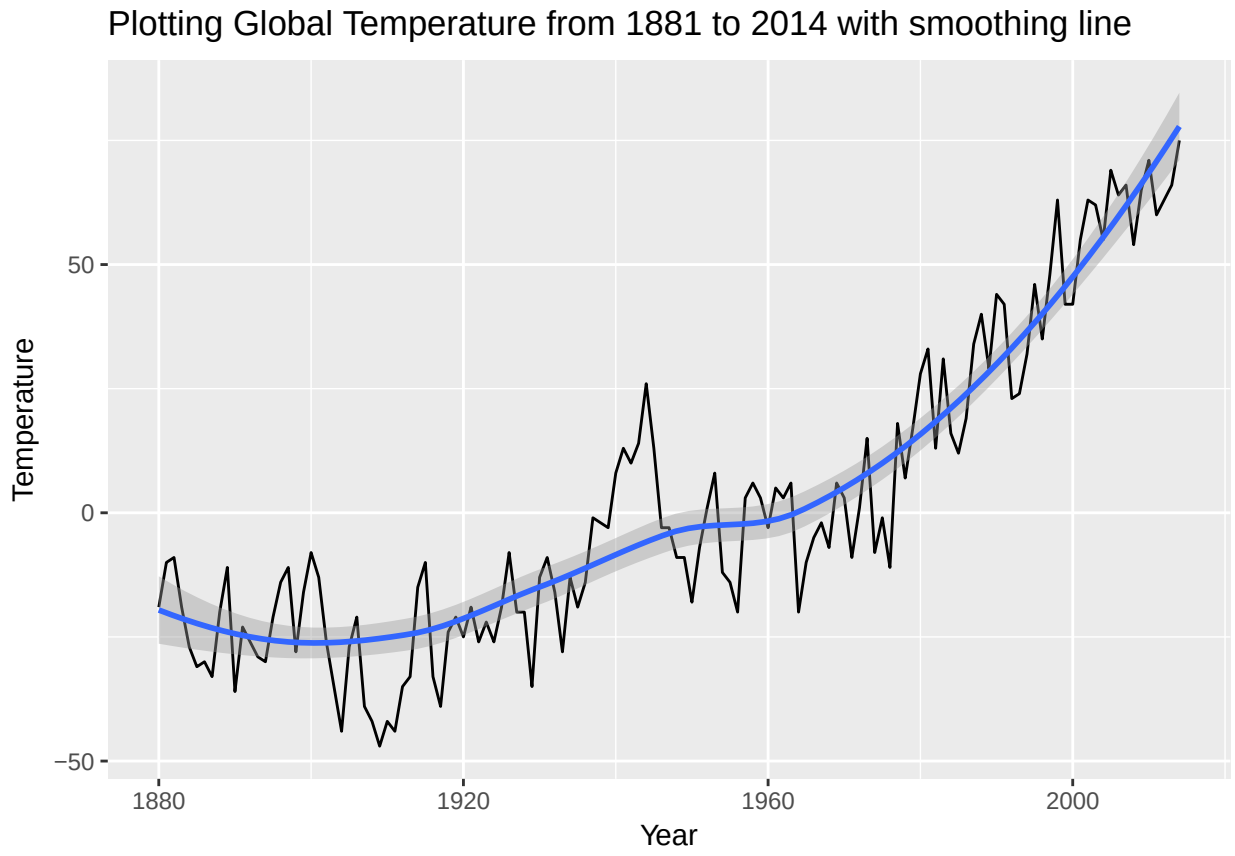
```
ggplot(RData2, aes(Year))+
  geom_line(aes(y=Glob))+
  ggtitle("Plotting Global Temperature Data from 1881 to 2014")+
  labs(y="Temperature")
```



Plotting global temperature with smooth line.

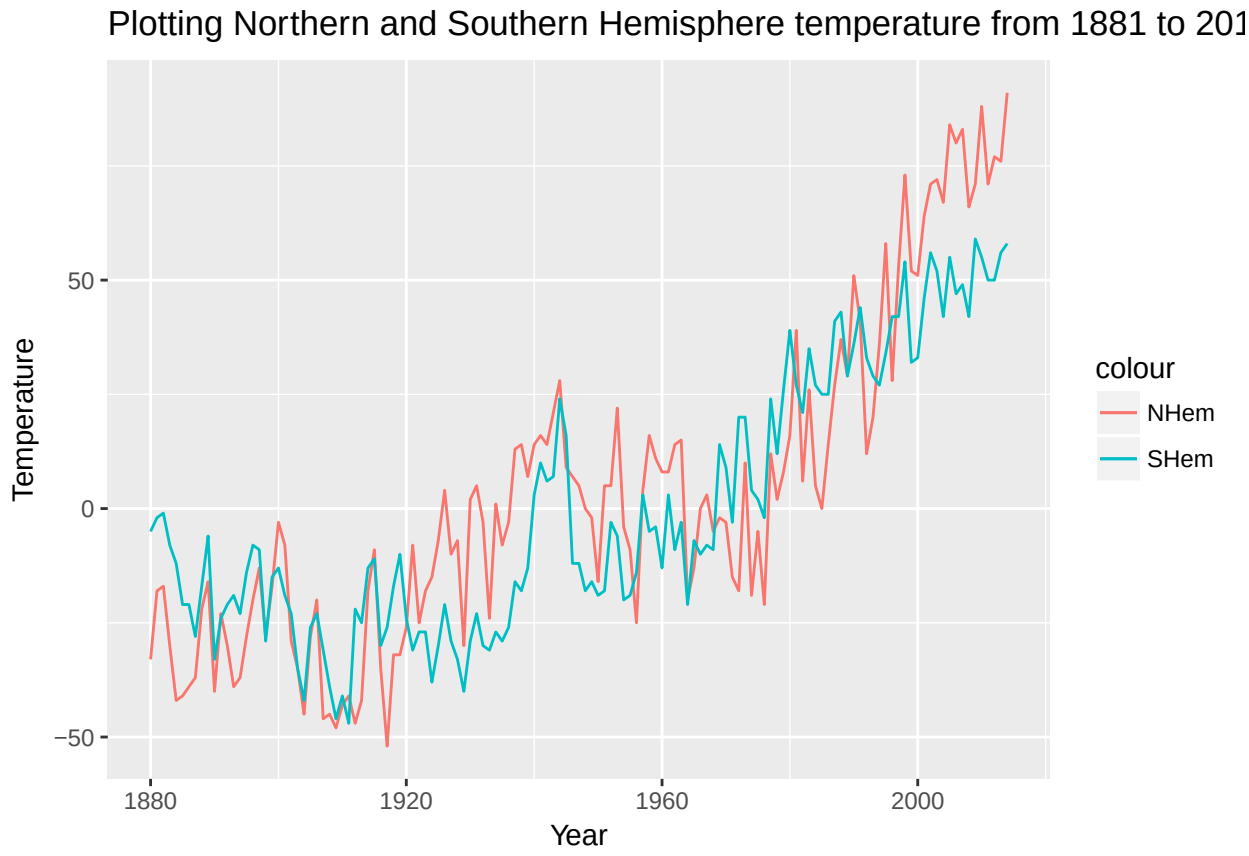
```
ggplot(RData2, aes(Year, Glob))+
  geom_line()+
  geom_smooth()+
  ggtitle("Plotting Global Temperature from 1881 to 2014 with smoothing line")+
  labs(y="Temperature")
```

```
## `geom_smooth()` using method = 'loess'
```



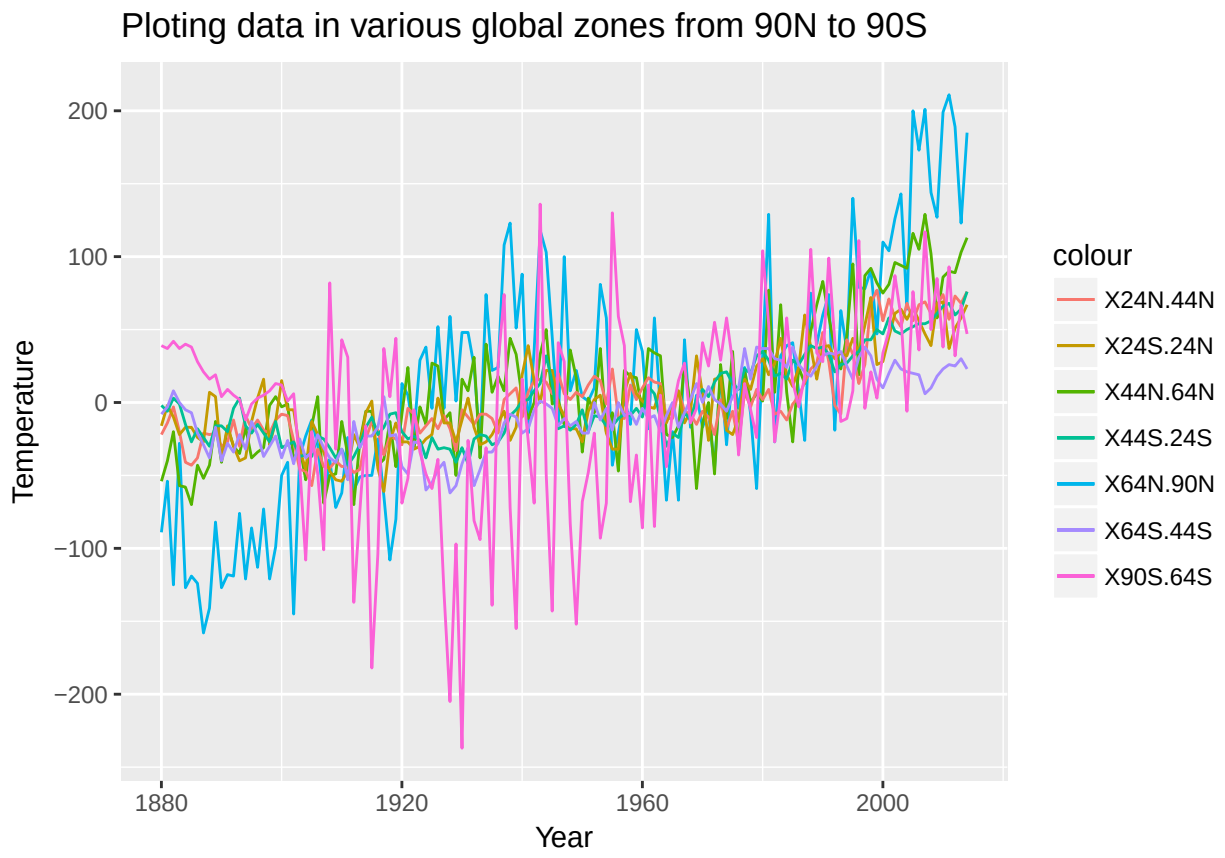
Plotting northern and southern hemisphere temperature.

```
ggplot(RData2, aes(Year))+
  geom_line(aes(y=NHem, colour="NHem"))+
  geom_line(aes(y=SHem, colour="SHem"))+
  ggtitle("Plotting Northern and Southern Hemisphere temperature from 1881 to 2014")+
  labs(y="Temperature")
```



Ploting data in various global zones from 90N to 90S.

```
ggplot(RData2, aes(Year))+
  geom_line(aes(y=X64N.90N, colour="X64N.90N"))+
  geom_line(aes(y=X44N.64N, colour="X44N.64N"))+
  geom_line(aes(y=X24N.44N, colour="X24N.44N"))+
  geom_line(aes(y=X24S.24N, colour="X24S.24N"))+
  geom_line(aes(y=X44S.24S, colour="X44S.24S"))+
  geom_line(aes(y=X64S.44S, colour="X64S.44S"))+
  geom_line(aes(y=X90S.64S, colour="X90S.64S"))+
  ggtitle("Ploting data in various global zones from 90N to 90S")+
  labs(y="Temperature")
```



Even in step 3, it is observed that all the temperature in various global zones has an increasing trend.

Conclusion

From both the different data sets and in all three-step analysis it is observed that over the period of time global temperature is increasing every year. That is why the entire world is concerned about global warming and is taking corrective measures to prevent it.