# Story Telling with Data - ISQA 8750 (Fall 2020)

# Assignment 1 :NASA GISS Temperature Analysis

# University of Nebraska at Omaha

# Professor : Sachin Pawaskar, Ph.D., MBA, MS

# Prepared By : Akhila Pillalamarri on 10/10/2020

***Background on the NASA GISS Temperature Analyis Project***

The GISS Surface Temperature Analysis (GISTEMP) is an estimate of global surface temperature change.From the NASA GISS website the history of the data are "Global temperature records start around 1880 because observations did not sufficiently cover enough of the planet prior to that time. Reference:(<https://www.giss.nasa.gov/>)

This Data set was analyzed in R, the report was written in R MarkDown file and compiled using Knit.

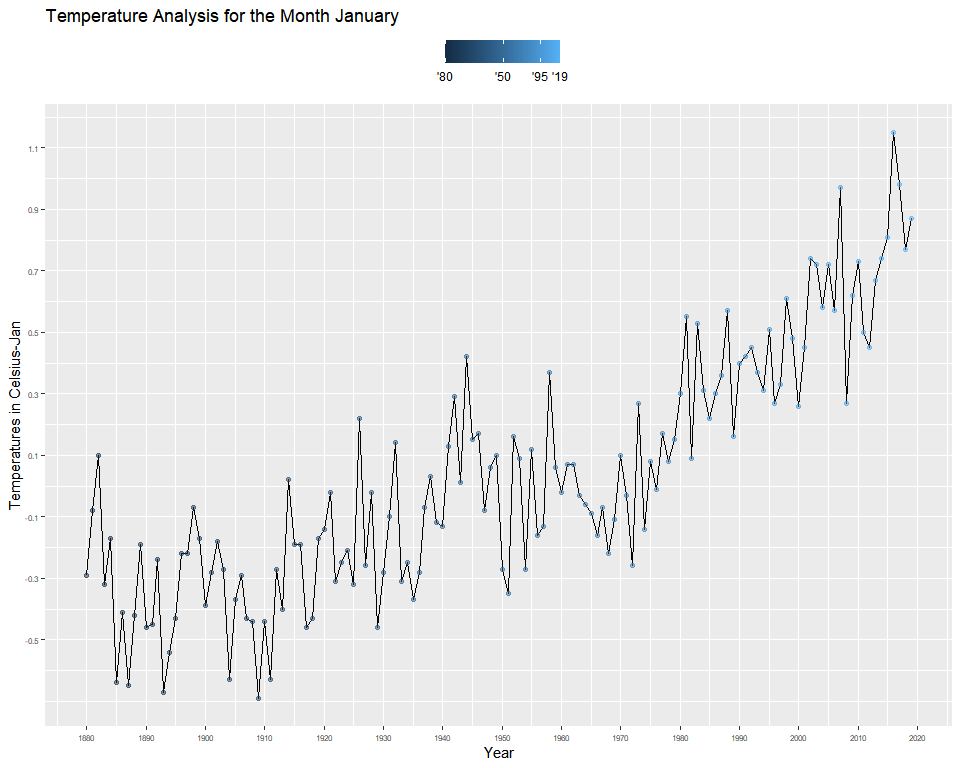
#The following libraries are useful in plotting maps and data manipulation:  
library(ggplot2)  
library(knitr)  
library(gridExtra)  
library(ggpubr)

library(readr)  
# Reading the csv file and assigning it into the climate1  
climate1 <- read.csv("NASA-GISTEMP-DataCSV.csv")  
#Reading another csv file and assigning it into climate2  
#NASA-GISTEMP-Data2CSV  
climate2 <- read.csv("NASA-GISTEMP-Data2CSV.csv")

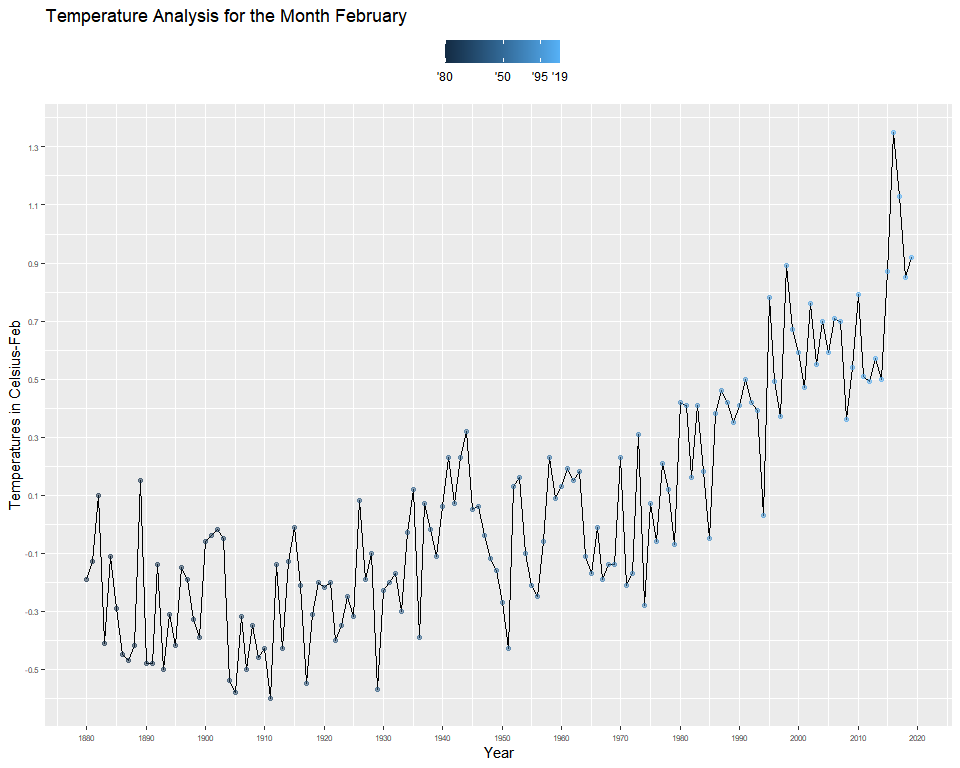
***Simple Line Plot| Temperature Analysis for different months***

Below is the simple line graph for different months from the year 1880 to 2019 which shows the temperature increase and with the help of color gradient it can be observed when the degrees were peaked.

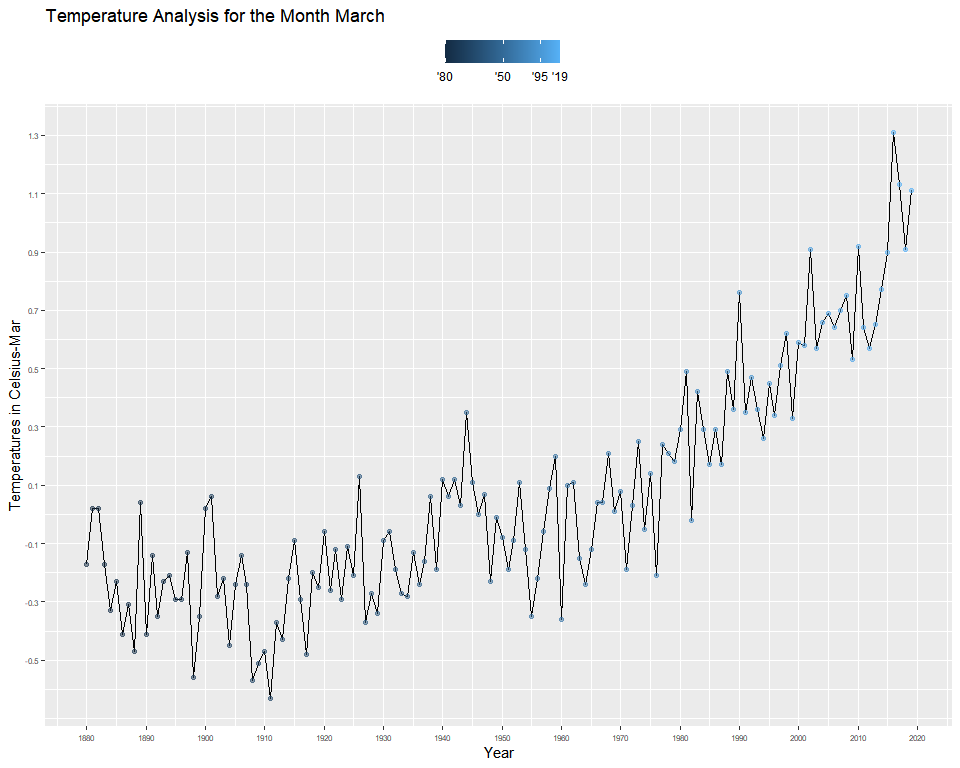
J1 <- ggplot(climate1, aes(x=Year,y=Jan/100))+geom\_line()+  
 theme(legend.position="top", axis.text=element\_text(size = 6))+  
 geom\_point( aes( color = Year), alpha = 0.5, size = 1.5)+  
 scale\_color\_continuous(name="", breaks = c(1880,1950, 1995, 2019),  
 labels = c("'80","'50","'95","'19"))+  
 ggtitle("Temperature Analysis for the Month January")+  
 xlab("Year")+ylab("Temperatures in Celsius-Jan")+  
 scale\_y\_continuous(breaks = seq(-.5, 1.5, by = .2))+  
 scale\_x\_continuous(breaks = seq(1880, 2020, by = 10))  
  
J1



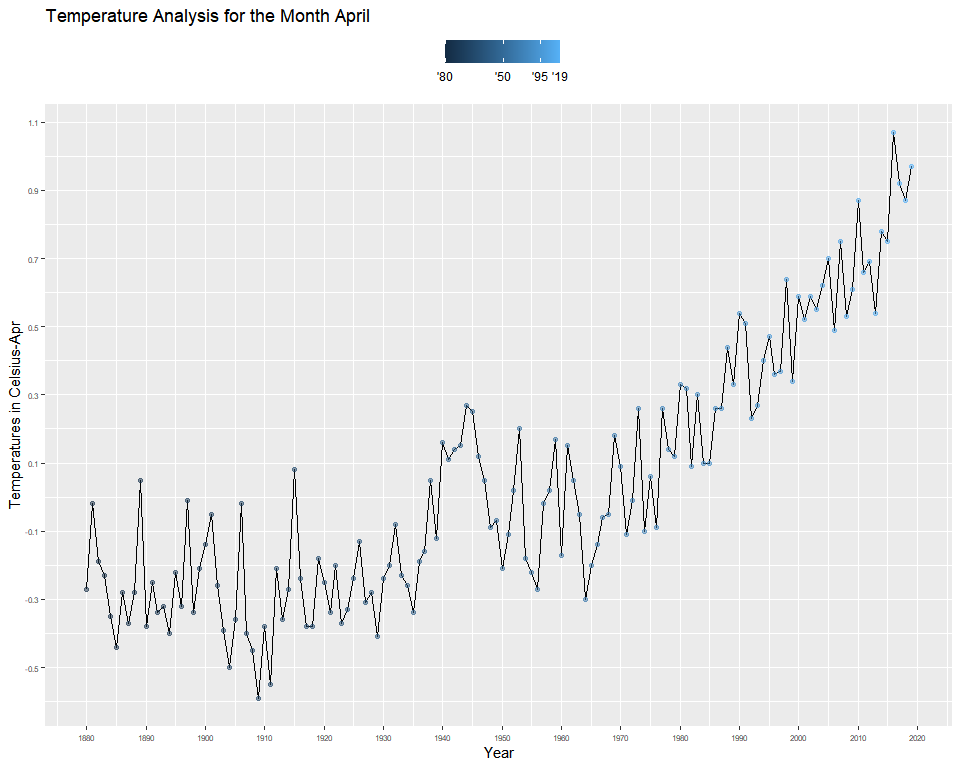
J2 <- ggplot(climate1, aes(x=Year,y=Feb/100))+geom\_line()+  
 theme(legend.position="top", axis.text=element\_text(size = 6))+  
 geom\_point( aes( color = Year), alpha = 0.5, size = 1.5)+  
 scale\_color\_continuous(name="", breaks = c(1880,1950, 1995, 2019),  
 labels = c("'80","'50","'95","'19"))+  
 ggtitle("Temperature Analysis for the Month February")+  
 xlab("Year")+ylab("Temperatures in Celsius-Feb")+  
 scale\_y\_continuous(breaks = seq(-.5, 1.5, by = .2))+  
 scale\_x\_continuous(breaks = seq(1880, 2020, by = 10))  
  
J2



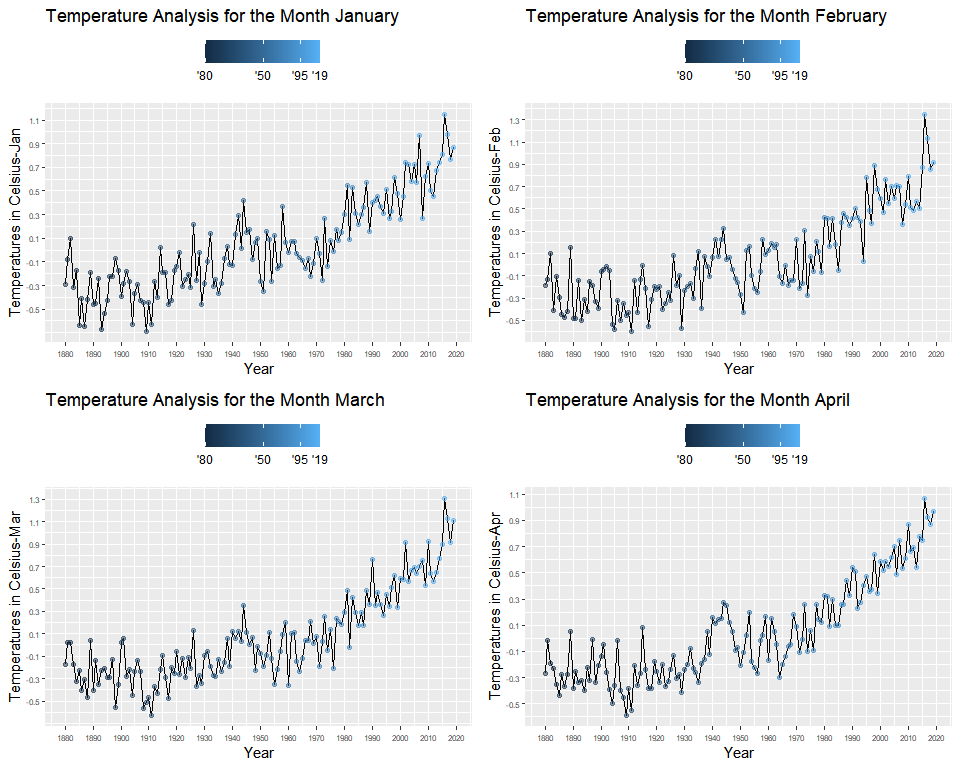
J3 <- ggplot(climate1, aes(x=Year,y=Mar/100))+geom\_line()+  
 theme(legend.position="top", axis.text=element\_text(size = 6))+  
 geom\_point( aes( color = Year), alpha = 0.5, size = 1.5)+  
 scale\_color\_continuous(name="", breaks = c(1880,1950, 1995, 2019),  
 labels = c("'80","'50","'95","'19"))+  
 ggtitle("Temperature Analysis for the Month March")+  
 xlab("Year")+ylab("Temperatures in Celsius-Mar")+  
 scale\_y\_continuous(breaks = seq(-.5, 1.5, by = .2))+  
 scale\_x\_continuous(breaks = seq(1880, 2020, by = 10))  
  
J3



J4 <- ggplot(climate1, aes(x=Year,y=Apr/100))+geom\_line()+  
 theme(legend.position="top", axis.text=element\_text(size = 6))+  
 geom\_point( aes( color = Year),alpha = 0.5, size = 1.5)+  
 scale\_color\_continuous(name="", breaks = c(1880,1950, 1995, 2019),  
 labels = c("'80","'50","'95","'19"))+  
 ggtitle("Temperature Analysis for the Month April")+  
 xlab("Year")+ylab("Temperatures in Celsius-Apr")+  
 scale\_y\_continuous(breaks = seq(-.5, 1.5, by = .2))+  
 scale\_x\_continuous(breaks = seq(1880, 2020, by = 10))  
  
J4



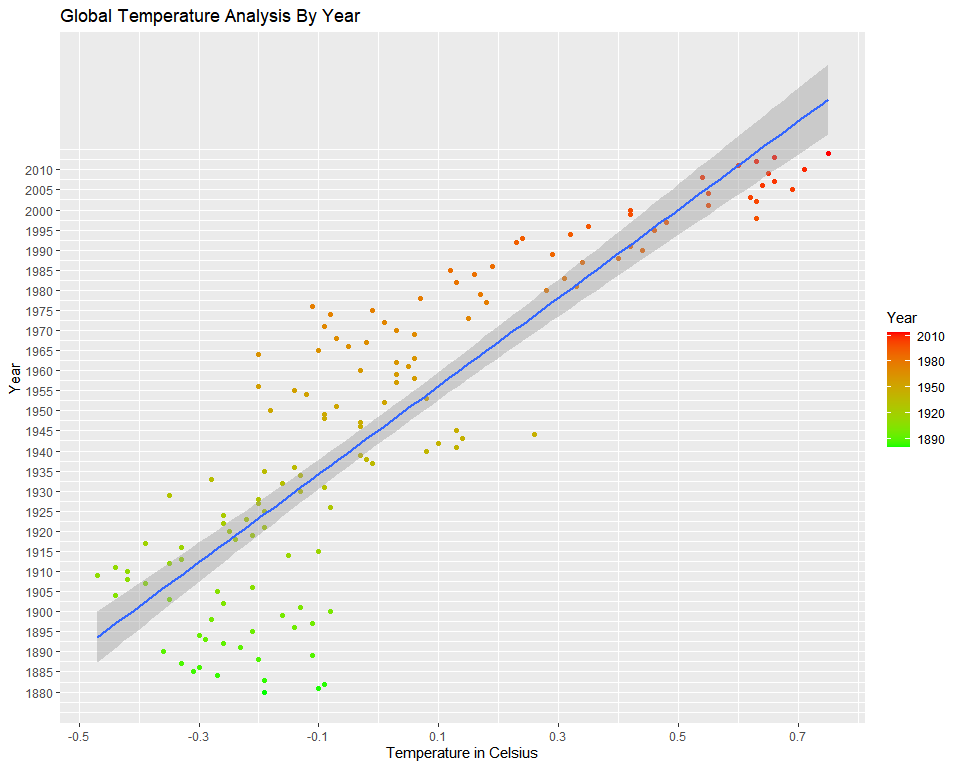
ggarrange(J1,J2,J3,J4,ncol=2,nrow=2)



#### **Global Temperature Analysis Using Scatter Graph**

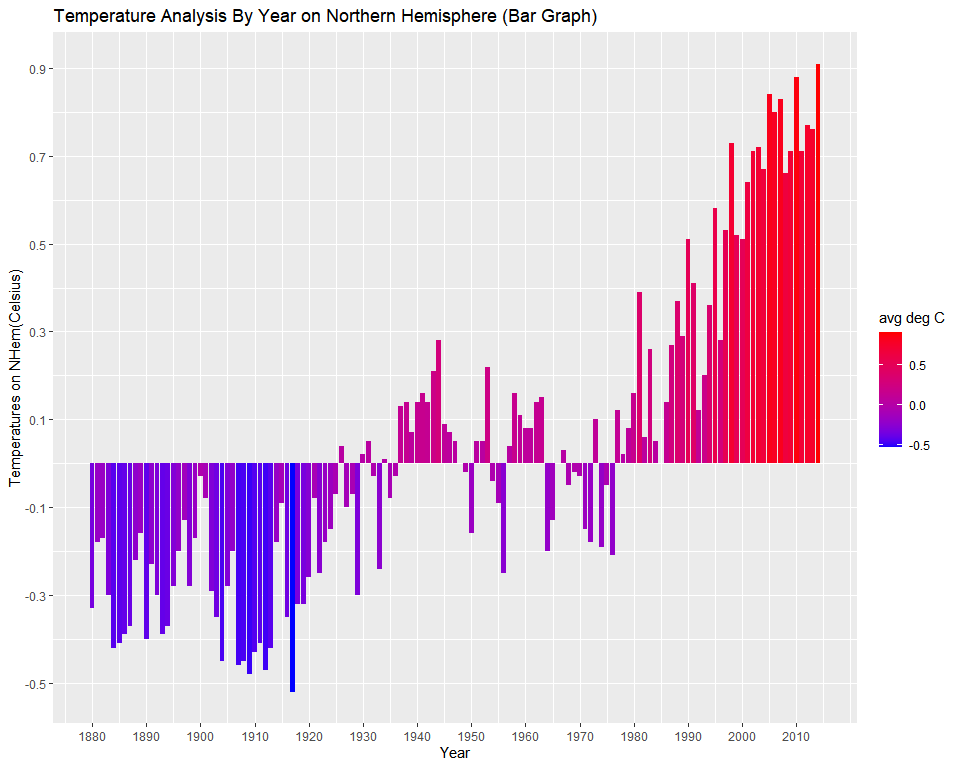
The following *scatter plot* is a *time series* plot of average absolute temperature anomalies.Here the graph is plotted between the year and the global temperatures(Celsius) which shows the temperature variations from year 1880 to 2019.Added linear model function “lm” for creating a simple regression model and color gradient is used to how the increase in temperatures year by year.

ggplot(climate2,aes(y= Year, x=Glob/100))+geom\_point(aes(color=Year))+  
 geom\_smooth(method="lm", formula=y ~ x, size=1)+  
 ggtitle("Global Temperature Analysis By Year")+  
 xlab("Temperature in Celsius")+ylab("Year")+  
 scale\_x\_continuous(breaks = seq(-.5, 1.5, by = .2))+  
 scale\_y\_continuous(breaks = seq(1880, 2014, by = 5))+  
 scale\_color\_gradient(low="green", high="Red")

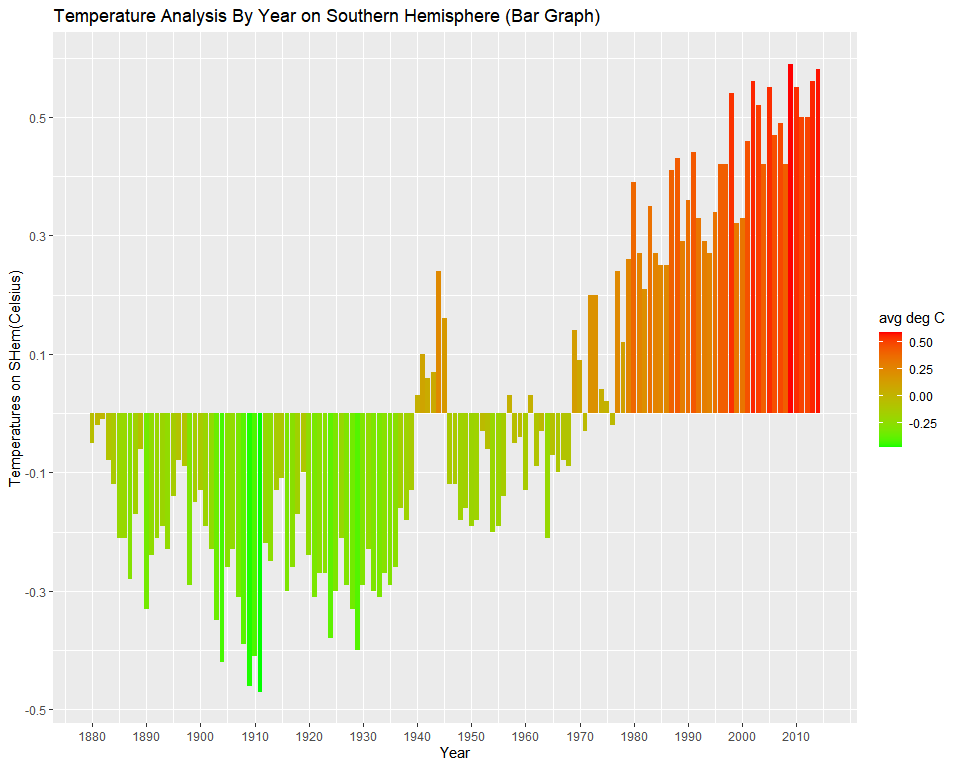


**Simple Bar Chart | NHem and SHem Temperature Anomalies** This is the bar chart interpretation of the Global Temperature Anomalies on the Northern Hemisphere and Southern Hemisphere.The color gradient is based on low to high. Created seperate graphs for both Northern Hemisphere and Southern Hemisphere and fially combined them using “gridextra” library to show both the graphs in one graph for easy comparison.

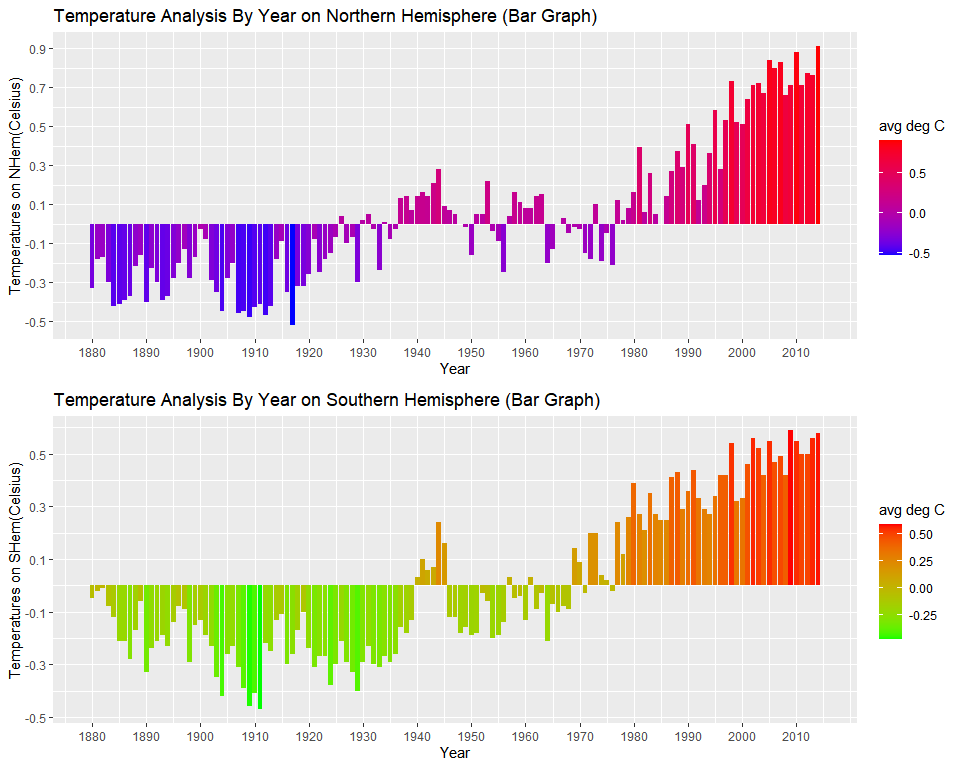
s1 <- ggplot(climate2,aes(x=Year, y=NHem/100))+  
 geom\_bar(stat="identity",aes(fill=NHem/100))+  
 ggtitle("Temperature Analysis By Year on Northern Hemisphere (Bar Graph)")+  
 xlab("Year")+ylab("Temperatures on NHem(Celsius)")+  
 scale\_y\_continuous(breaks = seq(-.5, 1.5, by = .2))+  
 scale\_x\_continuous(breaks = seq(1880, 2014, by = 10))+  
 scale\_fill\_gradient(low="blue", high="Red", name = "avg deg C")  
  
  
s1



#Bar Chart for SHem  
  
  
s2 <- ggplot(climate2,aes(x=Year, y=SHem/100))+  
 geom\_bar(stat="identity",aes(fill=SHem/100))+  
 ggtitle("Temperature Analysis By Year on Southern Hemisphere (Bar Graph)")+  
 xlab("Year")+ylab("Temperatures on SHem(Celsius)")+  
 scale\_y\_continuous(breaks = seq(-.5, 1.5, by = .2))+  
 scale\_x\_continuous(breaks = seq(1880, 2014, by = 10))+  
 scale\_fill\_gradient(low="Green", high="Red", name = "avg deg C")  
  
s2



#Combining both the NHem and SHem in one Graph  
  
  
grid.arrange(s1, s2)



***Temperature Analysis of months MAM, Northern and Southern Hemispheres*** Below is the line graph for the Temperature analysis of the months march, April and May also comparing with the temperatures in the Northern and Southern Hemispheres.

#comparing the data of MAM, NHem and SHem  
ggplot()+geom\_line(data=climate1,aes(x=Year, y=MAM/100,colour="green"))+  
 geom\_line(data=climate2, aes(x=Year, y=NHem/100, colour="blue")) +  
 geom\_line(data=climate2, aes(x=Year, y=SHem/100,colour="red"))+  
 ggtitle("Temperature Analysis of months MAM, Nothern, Southern Hemispheres")+  
 xlab("Year")+ylab("Temperatures in Celsius")+  
 scale\_y\_continuous(breaks = seq(-.5, 1.5, by = .2))+  
 scale\_x\_continuous(breaks = seq(1880, 2020, by = 10))

