NASA_GISS_Kyle_Hampton

Kyle Hampton

September 22, 2017

NASA GISS Surface Temperature Assignment Background

The Goddard Institute for Space Studies (GISS) at the National Aeronautics and Space Administration (NASA) publishes captured surface temperature data to identify changes in global surface temperature over time. NASA captures monthly surface temperature readings at various coordinate and hemisphere positions across the globe each year.

For additional background on the published surface temperature data NASA states the following, "We update the Goddard Institute for Space Studies (GISS) analysis of global surface temperature change, compare alternative analyses, and address questions about perception and reality of global warming. Satellite-observed nightlights are used to identify measurement stations located in extreme darkness and adjust temperature trends of urban and peri-urban stations for non-climatic factors, verifying that urban effects on analyzed global change are small." [1]

As recommended by NASA, a good portion of this analysis focuses on 12-month, average surface temperatures. Per NASA, "We suggest use of 12-month (and n×12) running mean temperature to fully remove the annual cycle and improve information content in temperature graphs." [1]

This assignment includes two data sets. The first data set captures annual and monthly global surface temperature data, including averages, for every year from 1880 to 2015. We will refer to this data set as the annual surface temperature data. The second data set captures annual global surface temperature data the northern and southern hemispheres as well as at specific latitude and longitude coordinates. We will refer to this data set as location-specific surface temperature data.

Assignment Plots & Visualizations

Plot: Annual Average Surface Temperature Since 1880

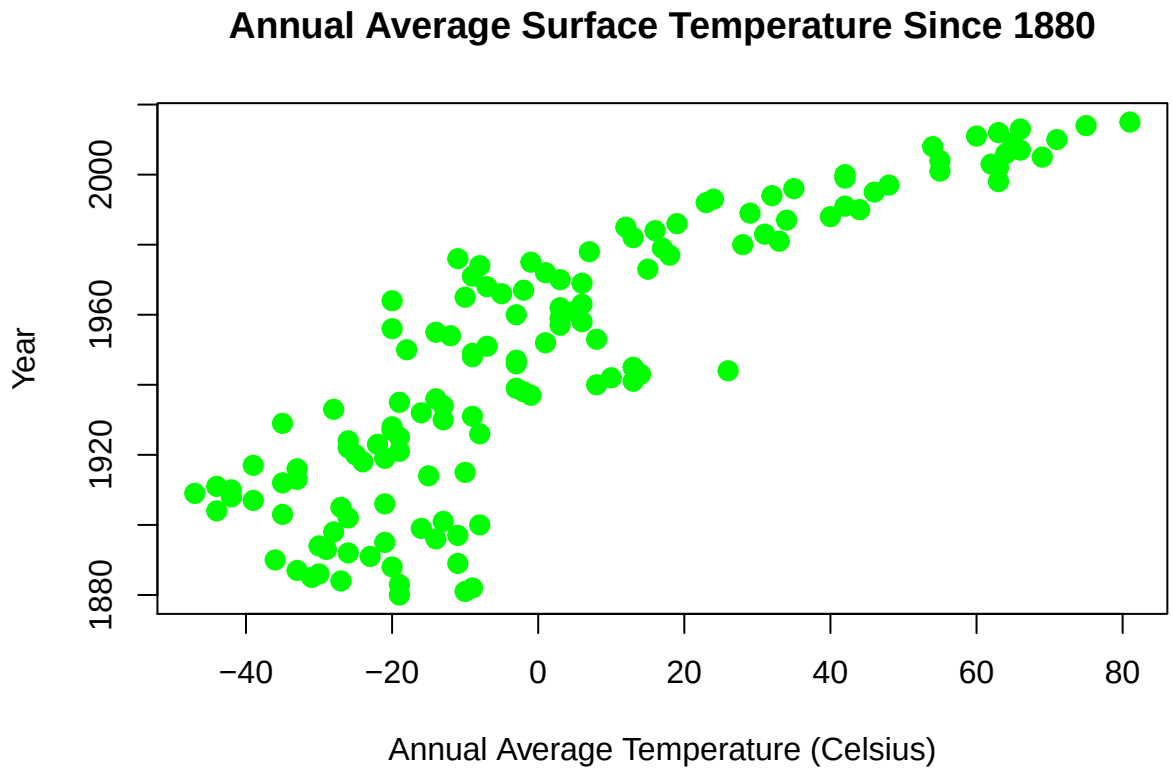
This first data plot displays global, annual average surface temperatures since 1880 with temperature plotted along the X-axis and years plotted along the Y-axis. This first visualization allows us to notice the data is not uniformly distributed. This visualization appears to indicate a relationship between temperature and years that shows temperature has increased since 1880, particularly starting around the 1930's and increasing significantly since then. We will continue to explore this relationship in subsequent visualizations.

```
library(readr)
nasa <- read_csv("C:/Users/kyleh/Google Drive UNO/UNO/Fall 2017/ISQA 8086/Assignment 1/NASA GISS Assignm
```

```
## Parsed with column specification:
## cols(
##   Year = col_integer(),
##   Jan = col_integer(),
##   Feb = col_integer(),
##   Mar = col_integer(),
##   Apr = col_integer(),
##   May = col_integer(),
##   Jun = col_integer(),
##   Jul = col_integer(),
```

```
##   Aug = col_integer(),
##   Sep = col_integer(),
##   Oct = col_integer(),
##   Nov = col_integer(),
##   Dec = col_integer(),
##   J.D = col_integer(),
##   D.N = col_integer(),
##   DJF = col_integer(),
##   MAM = col_integer(),
##   JJA = col_integer(),
##   SON = col_integer()
## )
```

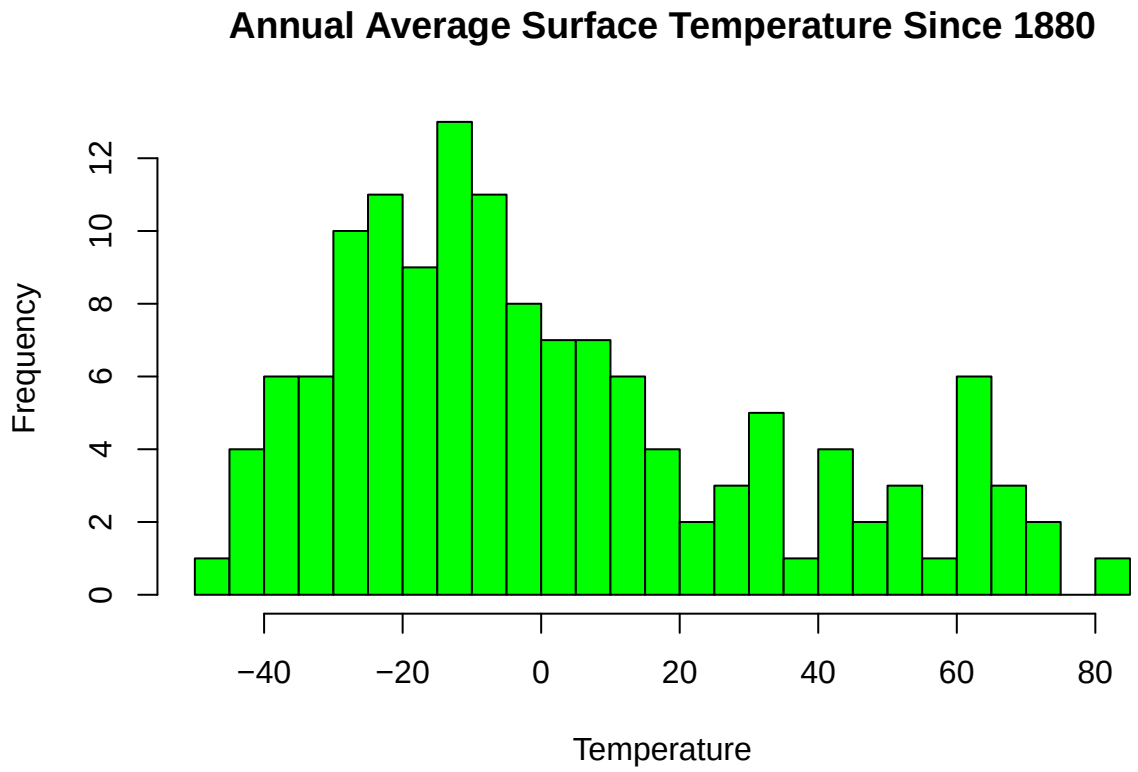
```
x <- nasa$J.D
y <- nasa$Year
data <- as.data.frame(cbind(x,y))
plot(data, xlab="Annual Average Temperature (Celsius)", ylab="Year", cex=2, col="green", pch=20, main="
```



Histogram: Annual Average Surface Temperature Since 1880

This second visualization takes the form of a histogram to give us different look at the same set of data in the first visualization. Immediately noticeable is that the histogram does not follow a standard, uniformly distributed bell curve. This histogram is very clearly skewed right, which further solidifies our earlier observation about increasing temperatures over the past century. If there were not changes in temperature, we would expect to see a standard bell curve; however, what we are looking at is a heavily skewed right bell curve of this histogram.

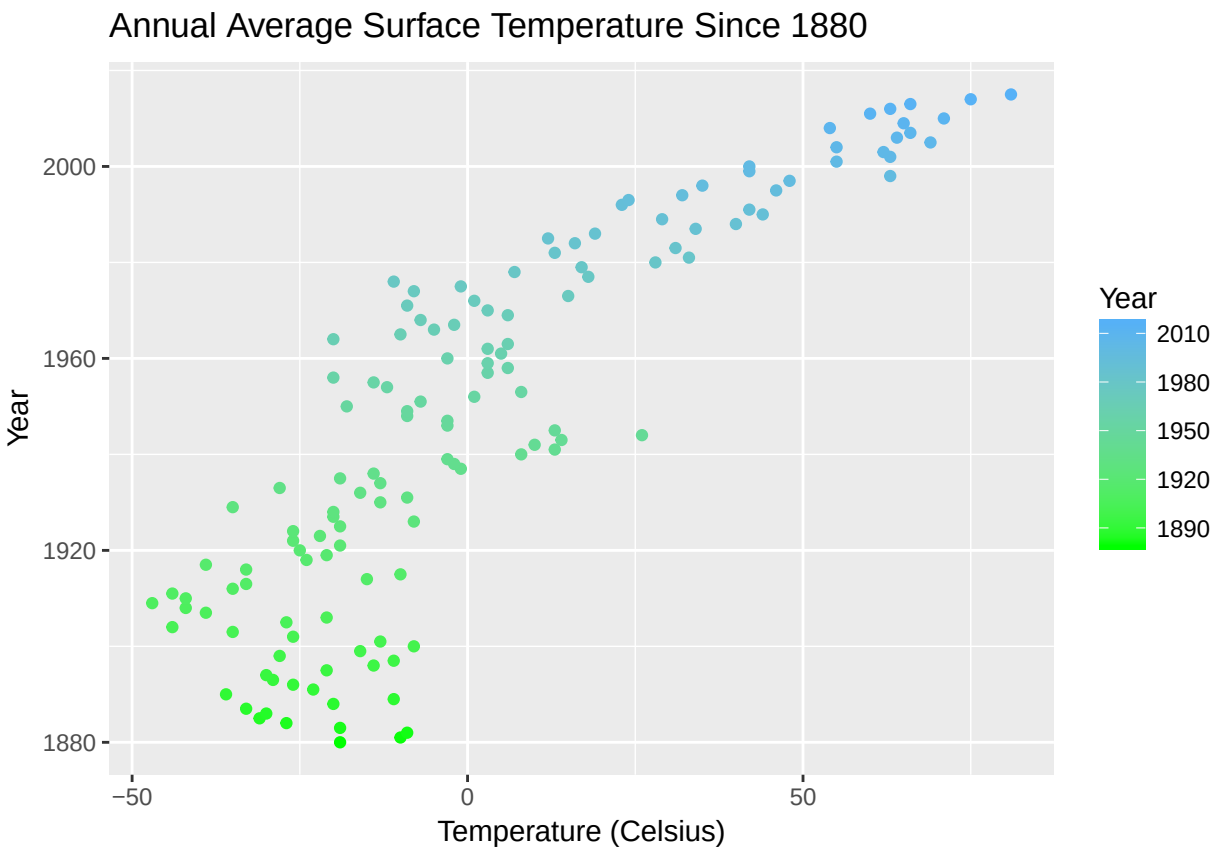
```
hist(nasa$J.D, breaks=20, main="Annual Average Surface Temperature Since 1880",xlab="Temperature", col=
```



GGPlot 1: Annual Average Surface Temperature Since 1880

Our third visualization utilizes GGPlot functionality to give us a similar, yet more detailed look at the same set of data as the prior two visualizations. This visualization utilizes a color gradient to easily distinguish data points by year. The key takeaway from this graph is that the temperature data over the past century is showing a significant increase. This visualization also appears to indicate that temperatures prior to the 1940's were relatively within the same range.

```
library(ggplot2)
nasaplot1 <- ggplot(nasa, aes(x=J.D, y=Year, color=Year)) + geom_point(aes(colour=Year)) + scale_color_
print(nasaplot1 + ggtitle("Annual Average Surface Temperature Since 1880") + labs(x="Temperature (Celsi
```

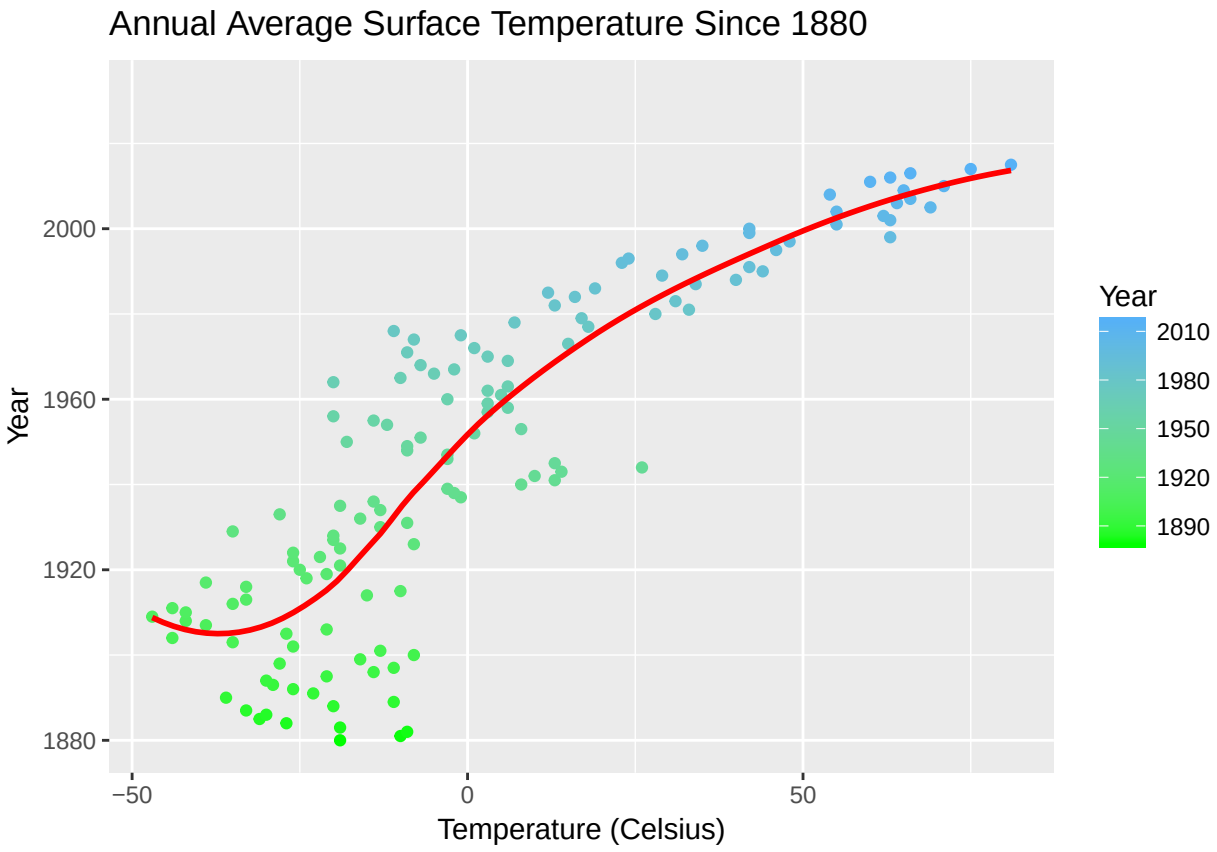


GGPlot 2: Annual Average Surface Temperature Since 1880 With Smoothing Curve

This next GGPlot is almost identical to the prior visualization; yet with the addition of a smoothing curve. This curve clearly highlights the change in temperature over the past century and also demonstrates that it appears the rate at which surface temperatures are increasing is decreasing, which is the key takeaway of this visualization.

```
library(ggplot2)
nasaplot1 <- ggplot(nasa, aes(x=J.D, y=Year, color=Year)) + geom_point(aes(colour=Year)) + scale_color_g
print(nasaplot1 + ggtitle("Annual Average Surface Temperature Since 1880") + labs(x="Temperature (Celsiu
```

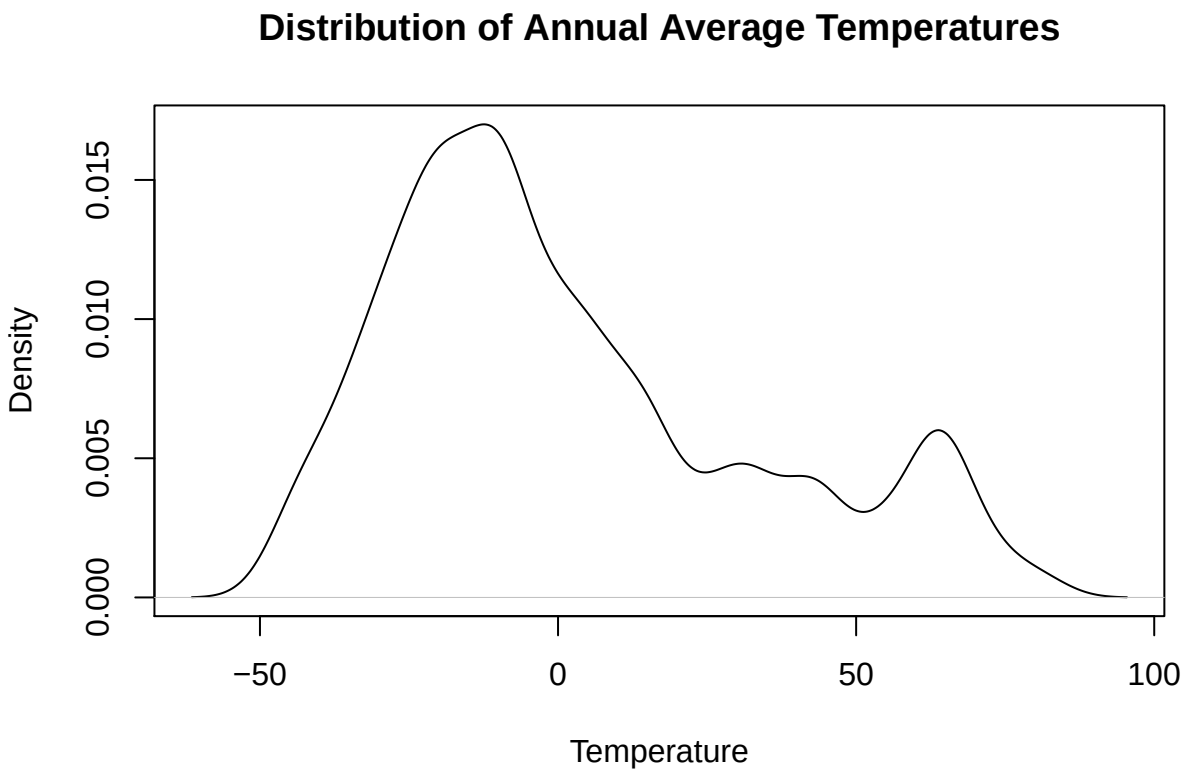
```
## `geom_smooth()` using method = 'loess'
```



Plots: Distribution of Annual Average Temperatures

The following two plots dive deeper into the distribution of annual average temperatures. Simply, this next visualization shows again that temperature is not uniformly distributed. The density of this chart leaves something to be desired as we can't exact dig into the structures of this visualization with this view.

```
plot(density((nasa$J.D),na.rm=TRUE,adjust=.5),main="Distribution of Annual Average Temperatures",xlab="
```



To look at this density distribution in a different way and dig deeper into the visualization, we can apply a logarithm to better understand the underlying structures of the visualization. This density distribution visualization further reinforces the notion of a non-standard relationship between temperature and time and further emphasizes the increasing temperatures in recent years.

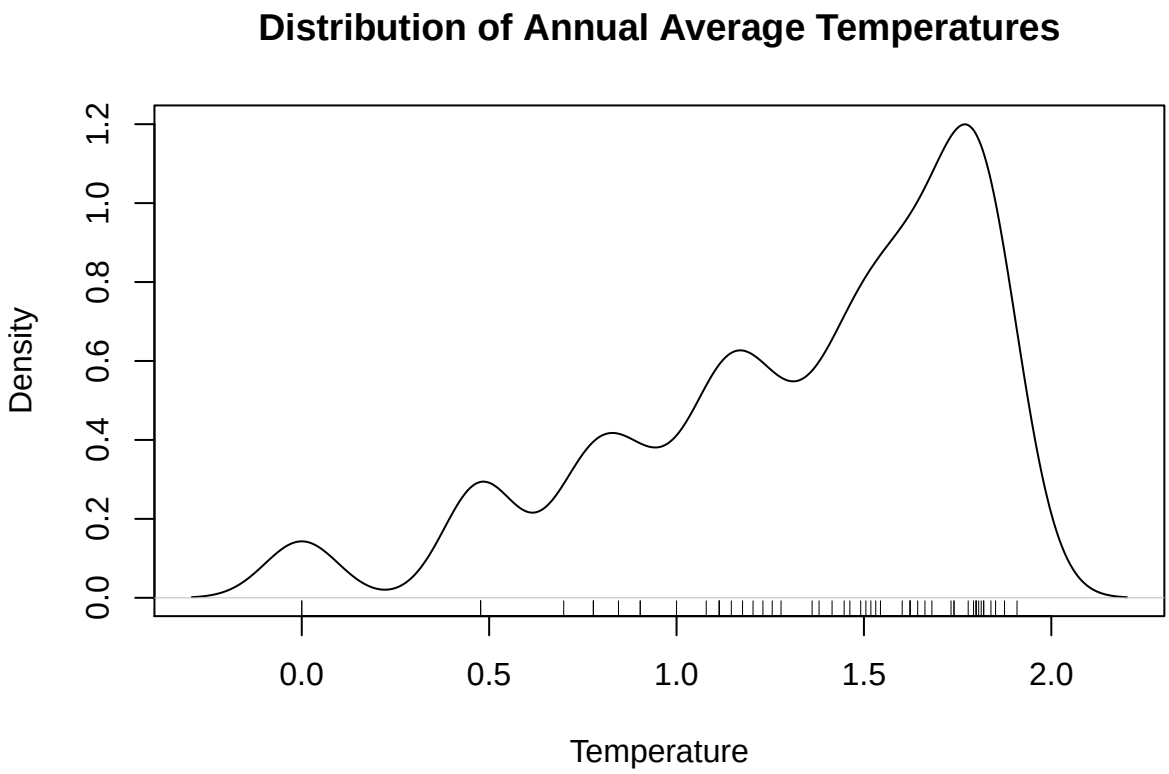
```
plot(density(log10(nasa$J.D),na.rm=TRUE,adjust=.5),main="Distribution of Annual Average Temperatures", x
```

```
## Warning in density(log10(nasa$J.D), na.rm = TRUE, adjust = 0.5): NaNs
## produced
```

```
rug(log10(nasa$J.D),na.rm=TRUE)
```

```
## Warning in as.vector(x): NaNs produced
```

```
## Warning in axis(side = side, at = at, labels = labels, ...): "na.rm" is not
## a graphical parameter
```



GGPlot: Hemisphere Annual Average Surface Temperatures Since 1880

This final visualization utilizes GGPlot to visualize northern and southern hemisphere average, annual temperatures since 1880. This plot demonstrates that the increase in temperature is not exclusive to any one hemisphere as the northern and southern hemispheres seem to oscillate back and forth. Additionally, this plot again shows the increase in surface temperatures in recent years. Also noteworthy, this visualization appears to show a growing divergence between northern and southern hemisphere temperatures and we should be curious to see if this difference continues in the coming years.

```
options(repos=structure(c(CRAN="https://cran.cnr.berkeley.edu/")))
nasa2 <- read_csv("C:/Users/kyleh/Google Drive UNO/UNO/Fall 2017/ISQA 8086/Assignment 1/NASA GISS Assig
```

```
## Parsed with column specification:
## cols(
##   Year = col_integer(),
##   Glob = col_integer(),
##   NHem = col_integer(),
##   SHem = col_integer(),
##   `24N-90N` = col_integer(),
##   `24S-24N` = col_integer(),
##   `90S-24S` = col_integer(),
##   `64N-90N` = col_integer(),
##   `44N-64N` = col_integer(),
##   `24N-44N` = col_integer(),
##   `EQU-24N` = col_integer(),
##   `24S-EQU` = col_integer(),
```

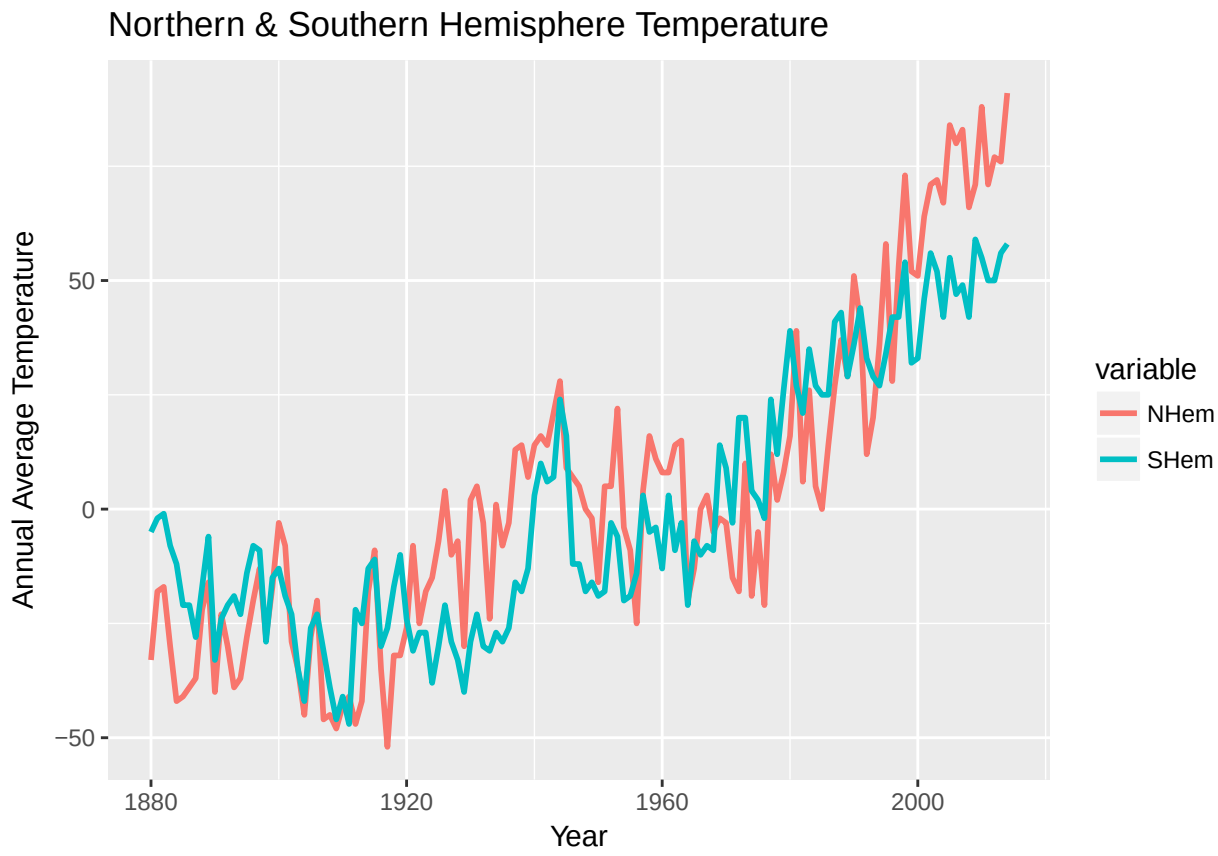
```
##   `44S-24S` = col_integer(),
##   `64S-44S` = col_integer(),
##   `90S-64S` = col_integer()
## )
```

```
install.packages("reshape")
```

```
## Installing package into 'C:/Users/kyleh/Documents/R/win-library/3.4'
## (as 'lib' is unspecified)
```

```
## package 'reshape' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\kyleh\AppData\Local\Temp\RtmpMla2X3\downloaded_packages
```

```
library(reshape)
a <- nasa2$Year
SHem <- nasa2$SHem
NHem <- nasa2$NHem
df <- data.frame(NHem,SHem,a)
df <- melt(df, id.vars='a')
ggplot(df, aes(x=a, y=value, fill=variable, colour=variable)) + geom_line(size=1) + xlab("Year") + ylab
```



Assignment Conclusions

This analysis of global surface temperatures since 1880 reveals a number key observations. First, since 1880, and in particular since 1930, global surface temperatures are increasing. Second, the rate at which global surface temperatures are increasing is decreasing as exemplified with the usage of a smoothing curve in our analysis. Third, there does not appear to be a difference in the rate at which northern and southern hemisphere temperatures are increasing except since the dawn of the new millennium. Since around the year 2000, the data demonstrates that the southern hemisphere temperature increases are decreasing at a much faster rate than the northern hemisphere.

With these key takeaways in mind, the overall conclusion from this data visualization analysis is that global surface temperatures since 1880 have increased significantly although temperatures have showed signs of subtly decreasing over recent decades.

References

[1] Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. Rev. Geophys., 48, RG4004, doi:10.1029/2010RG000345.