Case Study 2

Hang Tian
UB PERSON ID: 50413372

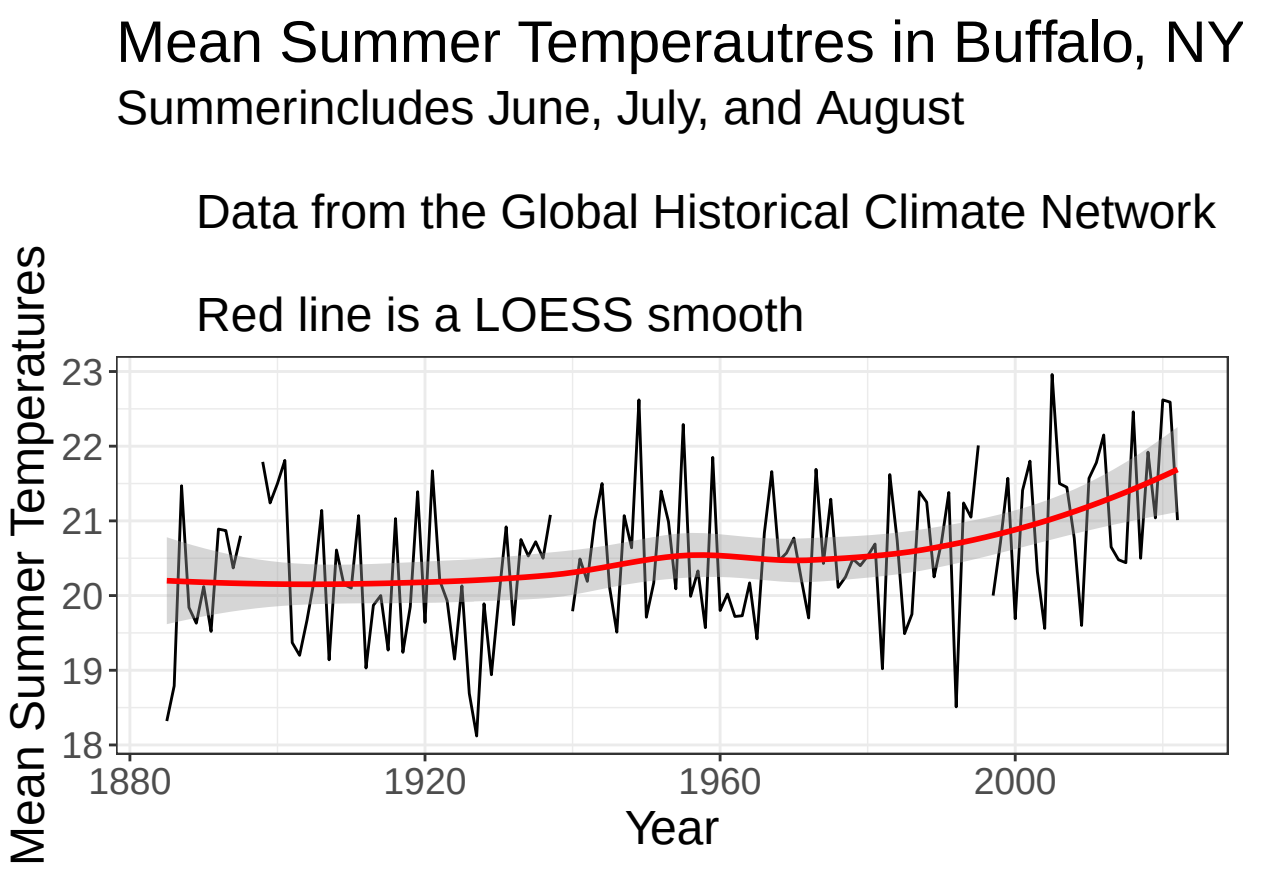
2022-09-08

1. Downloading data

```
library(ggplot2)
suppressPackageStartupMessages(library(tidyverse))
dataurl="https://data.giss.nasa.gov/tmp/gistemp/STATIONS/tmp_USW00014733_14_0_1/station.csv"
temp=read_csv(dataurl,
              skip=3, #skip the first line which has column names
              na="999.90", # tell R that 999.90 means missing in this dataset
              col_names = c("YEAR","JAN","FEB","MAR", # define column names
                            "APR","MAY","JUN","JUL",
                            "AUG","SEP","OCT","NOV",
                            "DEC","DJF","MAM","JJA",
                            "SON","metANN"),
              show_col_types = FALSE)
temp_summer=temp[c('YEAR','JJA')]
```

2. Make a flow chart

```
plt=ggplot(temp_summer,aes(YEAR,JJA))+
  geom_line()+
  theme_bw()+
  labs(title='Mean Summer Temperautres in Buffalo, NY',
      subtitle='Summerincludes June, July, and August\n
      Data from the Global Historical Climate Network\n
      Red line is a LOESS smooth',
      x='Year',
      y='Mean Summer Temperatures')+
  theme(title=element_text(size=18),axis.text=element_text(size=14))+
  geom_smooth(method='loess',formula='y~x',color='red')
suppressWarnings(print(plt))
```




```
ggsave(  filename = "Mean Summer Temperatures in Buffalo.png",
  width = 17,
  height = 9,
  units = "in",
  dpi = 300
)
```

```
## Warning: Removed 5 rows containing non-finite values (stat_smooth).
```