1 Introduction

1.1 Read Data

```
library(here)

## here() starts at /home/mwl04747/RTricks

library(xtable)

stations.active.oldest = read.csv(here("04_Regional_Climate_Trends", "Guides", "stations.act
```

1.2 Select and Evaluate State Data

```
stations.unique = unique(stations.active.oldest[,c("STATE", "STATE_NAME")])

xtab = xtable(stations.unique)
```

```
my.state = "CA"
```

2 Download Data from NOAA

2.1 Function to Download Data

```
# Select Stations in State
my.stations = subset(stations.active.oldest, STATE == my.state)

# Download Updated Station Data
i=1
here::here("04_Regional_Climate_Trends", my.stations$ID[i])

## [1] "/home/mwl04747/RTricks/04_Regional_Climate_Trends/USC00043157"

#station = data.frame(NULL)
for(i in 1:nrow(my.stations)){
  url = paste0("https://www.ncei.noaa.gov/pub/data/ghcn/daily/by_station/", my.stations$ID[i

print(i) # Print Index Number
download.file(url, paste0(here::here("04_Regional_Climate_Trends", "Data", "SP24/"), my.stat

assign(paste0("station", i), read.csv(gzfile(paste0(here::here("04_Regional_Climate_Trends",
```

```
# can't get the header named in loop!
#names(paste0("station",i)) <- c("ID", "DATE", "ELEMENT", "VALUE", "M-FLAG", "Q-FLAG", "S-FL

}
```

```
## [1] 1
## [1] 2
## [1] 3
## [1] 4
## [1] 5
```

```
names(station1) <- c("ID", "DATE", "ELEMENT", "VALUE", "M-FLAG", "Q-FLAG", "S-FLAG", "OBS-TI
names(station3) <- names(station2) <- names(station1)
names(station5) <- names(station4) <- names(station1)

# NAMES OF VARIABLES ARE INCORRECT??

  #ID = 11 character station identification code
  #YEAR/MONTH/DAY = 8 character date in YYYYMMDD format (e.g. 19860529 = May 29, 1986)
  #ELEMENT = 4 character indicator of element type
  #DATA VALUE = 5 character data value for ELEMENT
  #M-FLAG = 1 character Measurement Flag
  #Q-FLAG = 1 character Quality Flag
  #S-FLAG = 1 character Source Flag
  #OBS-TIME = 4-character time of observation in hour-minute format (i.e. 0700 =7:00 am); if
 #is available, the field is left empty
```

3 Process and Clean Data

3.1 Correct Values Units

```
station1$VALUE = station1$VALUE/10
```

3.2 Fix Dates!

```
#fix Dates
station1$Ymd = as.Date(as.character(station1$DATE), format = "%Y%m%d")
str(station1)
```

```
## 'data.frame': 224921 obs. of  9 variables:
##  $ ID      : chr  "USC00043157" "USC00043157" "USC00043157" "USC00043157" ...
```

```
##  $ DATE    : int  18670601 18670602 18670603 18670604 18670605 18670606 18670607 18670608
##  $ ELEMENT : chr  "PRCP" "PRCP" "PRCP" "PRCP" ...
##  $ VALUE   : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ M-FLAG  : chr  "" "" "" "" ...
##  $ Q-FLAG  : chr  "" "" "" "" ...
##  $ S-FLAG  : chr  "F" "F" "F" "F" ...
##  $ OBS-TIME: int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Ymd     : Date, format: "1867-06-01" "1867-06-02" ...
```

```
# Baseline Period 1961-1990

station1.baseline = subset(station1, Ymd >= "1961-01-01" & Ymd <= "1990-12-31")
a
```

```
## Error in eval(expr, envir, enclos):  object 'a' not found
```

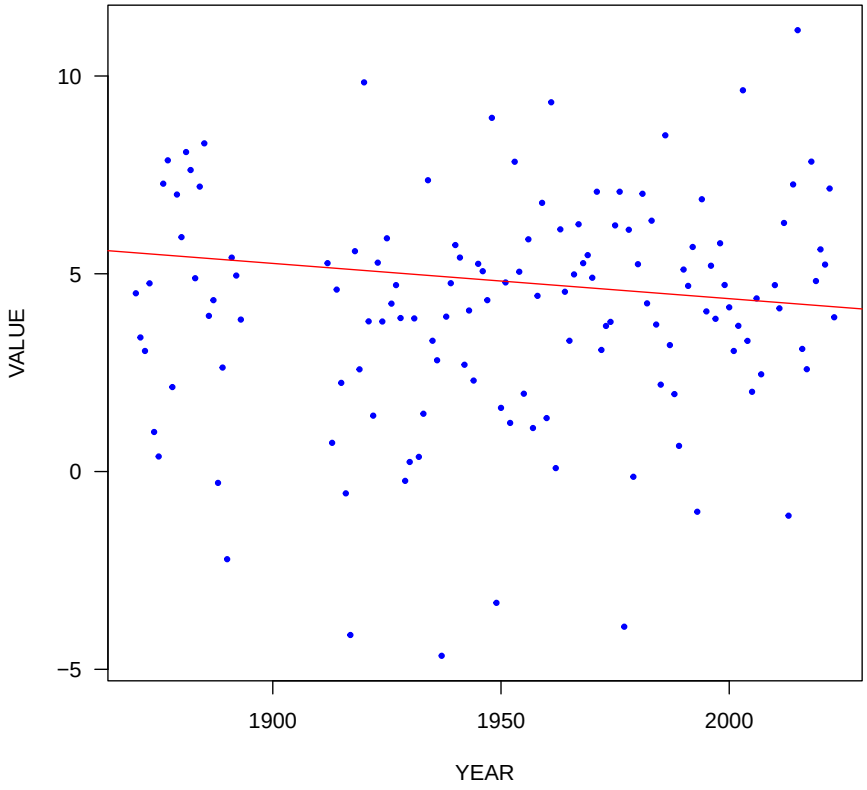
```
# Monthly Averages
station1$MONTH = as.numeric(format(station1$Ymd, "%m"))
station1$YEAR = as.numeric(format(station1$Ymd, "%Y"))
station1.monthly = aggregate(VALUE ~ MONTH + YEAR,
                             data = subset(station1, ELEMENT == "TMAX"), mean)
str(station1.monthly)
```

```
## 'data.frame': 1601 obs. of  3 variables:
##  $ MONTH: num  1 2 3 4 5 6 7 8 9 10 ...
##  $ YEAR : num  1870 1870 1870 1870 1870 1870 1870 1870 1870 1870 ...
##  $ VALUE: num  4.51 7.02 7.14 16.06 20.64 ...
```

```
# lm
for(i in unique(station1.monthly$MONTH)){
  temp.lm <- lm(VALUE ~ YEAR, data = subset(station1.monthly, MONTH==i))

}

plot(VALUE ~ YEAR, data = subset(station1.monthly, MONTH == 1),
     las=1, pch=19, col = "blue", cex=.5)
abline(coef(temp.lm), col = "red")
```


10
5
0
−5
VALUE
1900
1950
2000
YEAR