2.3 Cleaning Raw Climate Data

Manual of Applied Spatial Ecology

3/11/2022

1. Exercise 2.3 - Download and extract zip folder into your preferred location

2. Set working directory to the extracted folder in R under Session - Set Working Directory...

3. Now open the script "Read_FilesScript.Rmd" and run code directly from the script

4. No packages are needed for this exercise, these are base R functions

5. For each csv file, save as Excel Worsheet 1997-2003 if importing to NCSS or keep in csv or txt if using R.

6. Take out all non-Jan months for every year

7. Files will need to meet the following criteria but will be addressed in Exercise 2.4: Each weather station must have records for at least 10 of the 11 Januaries. Each weather station must have at least 95% of daily records for those Januaries. This means at least 325 days for 11 seasons and 295 days for 10 seasons.

Snow Depth (SNWD) 66 stations Maximum temp (TMAX) 69 stations Minimum temp (TMIN) 68 stations

8. The code that follows should have all files in the same folder but not the R script or any R files or code will not run. The code below brings in each text file and summarizes the data for each weather station as instructed in the code.

```
# Vector of files names in working directory
files <- list.files(pattern = ".txt")

# Total number of files in working directory (for loop below)
n.files <- length(files)

# Container to hold text files
files.list <- list()

#Populate the container files.list with climate data sets
files.list <- lapply(files, read.table, header =T, sep="\t")

#Set up matrix for weather station summary data
m1 <- matrix(NA,ncol=8,nrow=n.files)

#Loop for running through all weather station files
for(i in 1:n.files){

    # Assign elevation
        m1[i,1] <- files.list[[i]][1,10]

    #Assign Lat
        m1[i,2] <- files.list[[i]][1,11]

    #Assign Long
```

```
        m1[i,3] <- files.list[[i]][1,12]

    #Calculate mean snow depth
        SNWD_mm <- mean(files.list[[i]][,7],na.rm=T)

    #Convert snow depth mean to inches
    SNWD_in <- SNWD_mm/25.4

    #Assign snow depth
    m1[i,4] <- SNWD_in

    #Calculate mean maximum temp
        TMAX_C <- mean(files.list[[i]][,8],na.rm=T)

    #Convert max temp to F
    TMAX_F <- TMAX_C*0.18 + 32

    #Assign max temp
    m1[i,5] <- TMAX_F

    #Calculate mean minimum temp
    TMIN_C <- mean(files.list[[i]][,9],na.rm=T)

    #Convert min temp to F
    TMIN_F <- TMIN_C*0.18 + 32

    #Assign min temp
    m1[i,6] <- TMIN_F

    #Reassign GHCN number
    GHCN <- toString(files.list[[i]][1,1])

    #Assign Station Name
    m1[i,7] <- GHCN

    #Reassign Station Name
    SN <- toString(files.list[[i]][1,2])

    #Assign Station Name
    m1[i,8] <- SN
}

colnames(m1) <- c("Elevation","Lat","Long","SNWD","TMAX","TMIN","GHCN","Station")
write.csv(m1,paste(".","\\output.csv",sep=""))

#Removes quotation marks in output table
m1 <-noquote(m1)
m1[1:5,]
```

```
##      Elevation Lat      Long      SNWD              TMAX
## [1,] 520       42.249   -77.758   15.5581287633164 31.9693841642229
## [2,] 457.2     42.1     -78.749   7.69803623944597 30.1751125401929
## [3,] 452       42.303   -78.018   5.2936014962939  30.8728823529412
## [4,] 341.4     42.348   -77.347   4.12277728528967 32.2475747508306
```

```
## [5,] 80.2      40.83333 -75.08333 NaN             36.4575221238938
##      TMIN             GHCN        Station
## [1,] 13.2995307917889 USC00300085 ALFRED
## [2,] 14.4662580645161 USC00300093 ALLEGANYSP
## [3,] 11.6874486803519 USC00300183 ANGELICA
## [4,] 13.5491029900332 USC00300448 BATH
## [5,] 19.6215384615385 USC00280734 BELVIDERE
```