pollutantmean.R

markdeegan

Thu Aug 13 12:16:11 2015

```
pollutantmean <- function(directory, pollutant, id = 1:332) {
    ## 'directory' is a character vector of length 1 indicating
    ## the location of the CSV files

    ## 'id' is an integer vector indicating the monitor ID numbers
    ## to be used

    print(paste("We will be reading",length(id),"files from the directory:",directory));

    ## 'pollutant' is a character vector of length 1 indicating
    ## the name of the pollutant for which we will calculate the
    ## mean; either "sulfate" or "nitrate".
    if((pollutant == "nitrate") || (pollutant == "sulfate"))
    {
      print(paste(pollutant, ":is a valid pollutant name, so we will continue"));
    }
    else
    {
      # print(paste(pollutant, " is not a valid pollutant name, so we will exit here."));
      stop(paste(pollutant, "is not a valid pollutant name, so we will exit here."));
    }


    ## Return the mean of the pollutant across all monitors list
    ## in the 'id' vector (ignoring NA values)
    ## NOTE: Do not round the result!

    # add all data from the specified .csv files to the allData data frame
    # this code loops through all the values in the id argument to the function call
    # and for each, either creates (if it is the first value) or appends the records read from
    # the .csv file to the allData data frame
    for( i in id) # for each numeric in the id list provided (or not provided) as an argument
    { # begin for loop

      # this converts single or double-digit file numbers into a three-digit file number
      # padded with leading 0
      # for example 1 becomes 001, 21 becomes 021 and 321 remains 321
      fileNum <- formatC(i, width=3,flag="0");

      # this builds the file name using the directory, the padded file number and the .csv extension
      fname <- paste(directory,"/",fileNum,".csv",sep="");
      print(fname);
      # this checks if the allData frame already exists,
      # if not, then we create it by reading in a .csv file
      if(!exists("allData")) {
        allData <- read.csv(fname, header=TRUE);
      } # end if statement
```

```
      # if it does exist, then we add the data from the data file to the excisting data frame using rbi
      else {
        allData <- rbind(allData, read.csv(fname, header=TRUE));
      } # end else statement
    } # end for loop

    nrow(allData);

    good <- na.omit(allData[pollutant]);
    mean(good[2,])
}
```