Homework_11

Megan O’Connor

2024-04-28

NOTE: The code that is uncommented at the end is adjusted so that the homework will knit, however the code that is commented is the code I worked on in class. When double checking with my classmates (we have nearly exactly the same code besides names of our functions, directories, etc. my directories would not change, and after spell checking my directories, etc, I’m still unsure the issues with it. That work is displayed first.

1. Setting Up BarracudaR

```
library(log4r)
```

```
## Warning: package ’log4r’ was built under R version 4.1.2
```

```
##
## Attaching package: ’log4r’
```

```
## The following object is masked from ’package:base’:
##
##     debug
```

```
library(TeachingDemos)
library(tidyverse)
```

```
## Warning: package ’tidyverse’ was built under R version 4.1.2
```

```
## Warning: package ’tibble’ was built under R version 4.1.2
```

```
## Warning: package ’tidyr’ was built under R version 4.1.2
```

```
## Warning: package ’readr’ was built under R version 4.1.2
```

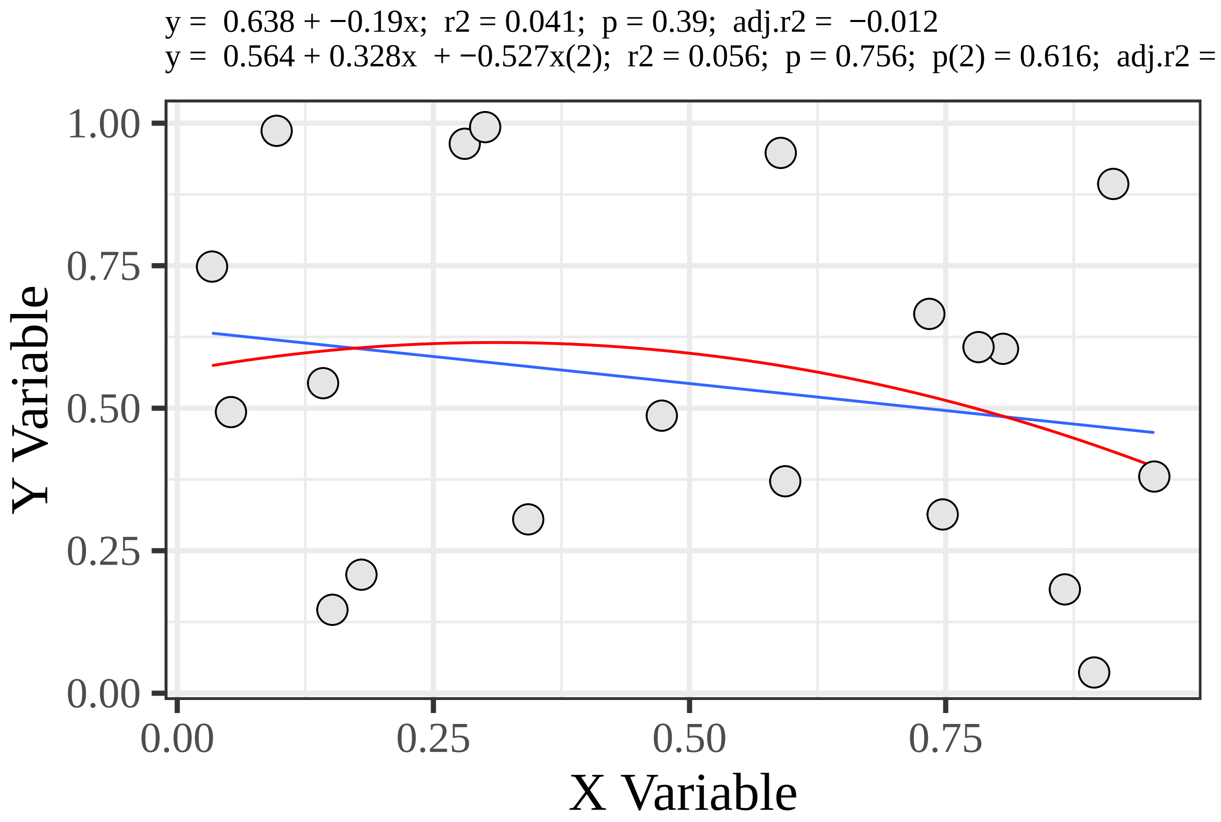
```
## Warning: package ’stringr’ was built under R version 4.1.2
## Warning: package 'forcats' was built under R version 4.1.2

## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.3     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(pracma)
```

```
##
## Attaching package: 'pracma'
##
## The following object is masked from 'package:purrr':
##
##     cross
```

```r
library(ggmosaic)

# source function files ----

setwd("~/Desktop/O'Connor_Biol6100/Homework11/barracudar")

source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/DataTableTemplate.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/AddFolder.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/BuildFunction.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/MetaDataTemplate.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/CreatePaddedLabel.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/InitiateSeed.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/SetUpLog.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/SourceBatch.R")

source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/QBox.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/QCon1.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/QCon2.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/QHist.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/QLogis.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/QScat.R")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



```
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/QBub.R")
source("~/Desktop/O'Connor_Biol6100/Homework11/barracudar/QContour.R")


library(stringr)


setwd("~/Desktop/O'Connor_Biol6100/Homework11/NEON_Data")
```

2.

```
# # getting our file names using list.files
#
# fileslists <- list.files("~/Desktop/O'Connor_Biol6100/Homework11/NEON_Data",pattern="BART")
#
# # use a for loop to pull out files
# # paste() or paste0() function puts strings together
# # paste0("Here is ","the ","filepath: ", filelist[1]) # example
#
# # make an empty vector
# file_names <- c()
#
# # make matrix of file names
# for (i in 1:10) {
#   setwd(paste0("~/Desktop/O'Connor_Biol6100/Homework11/NEON_Data","/", fileslists[i])) # create new f
#
#   file_names[i] <- list.files(pattern="countdata") # any time the file has "countdata" in it, it pull
# }
#
# file_names
```

Questions 3, 4, 5

```
# # - empty/missing cases
# # - extract the year from the file name
# # - total number of individuals
# # - number of unique species found
#
#
# information_func <- function(fileslists,file_names) {
#
#   metadata <- matrix(0,length(fileslists),4)
#   colnames(metadata) <- c("File","Year","total_indivs", "spec_rich")
#
#   for (i in 1:10) { # one csv per folder
#
# # which folder to go into
# setwd(paste0("~/Desktop/O'Connor_Biol6100/Homework11/NEON_Data","/", fileslists[i]))
# # # csv data
# csv_data <- read.csv(file_names[i])
#
# # reading in the data
# csv_data
# dfcsv <- data.frame(csvdata)
# dfcsv
#
# # getting all the categories now
#
# # empty/missing cases
# dfcsv <- na.omit(dfcsv[,1:20])
# dfcsv
#
# # year
# year <- str_sub(file_names[i],43,46)
#
# # total individuals
# total_individuals <- nrow(dfcsv)
#
# # total unique species
# species_richness <- length(unique(dfcsv[,12]))
#
#  # fill it in
# metadata[i,1:4] <- c(filelist[i], year, total_individuals, species_richness)
# # number of items to replace is not a multiple of replacement length
```

```r
#
#   }
#
# return(metadata)
#
#   }
#
# information_func(fileslists,file_names)
```

ADJUSTED CODE

```r
setwd("~/Desktop/O'Connor_Biol6100/Homework11/NEON_Data")

# Retrieve file lists with pattern "BART"
fileslists <- list.files(pattern = "BART")

# Initialize file_names vector
file_names <- character(length(fileslists))

# Loop to populate file_names with full file paths
for (i in seq_along(fileslists)) {
  # Construct full file path
  file_path <- file.path("~/Desktop/O'Connor_Biol6100/Homework11/NEON_Data", fileslists[i], "countdata")

  # Check if file exists
  if (length(dir(file_path)) > 0) {
    file_names[i] <- dir(file_path)
  }
}

# Remove empty elements from file_names
file_names <- file_names[file_names != ""]

# Information function
information_func <- function(fileslists, file_names) {
  metadata <- matrix(0, nrow = length(fileslists), ncol = 4)
  colnames(metadata) <- c("File", "Year", "Total Indivs", "Spec Rich")

  for (i in seq_along(fileslists)) {
    # Read CSV file
    csv_data <- read.csv(file_names[i])

    # Extract year from file name
    year <- as.numeric(substr(fileslists[i], 43, 46))

    # Total individuals
    total_individuals <- nrow(csv_data)

    # Unique species richness
    species_richness <- length(unique(csv_data[, 12]))
    # Fill metadata matrix
    metadata[i, ] <- c(fileslists[i], year, total_individuals, species_richness)
  }

  return(metadata)
}

# information_func(fileslists, file_names)
```