Assignment 3 Rankall Function (getting the ranked hospitals incorrectly)

code × dataframe × output × assignment3 × rankall × + Add Tag

Manuel Alejandro García Acosta · 10 days ago

Hi everyone, I'm trying to accomplish Assignment 3 and I have trouble with the rankall function. So far I'have managed to create a list with de apply functions that has my ordered data frames. Nevertheless I can't figure out how to substract correctly the names of the hospitals and bind them with the name of their respective states

```
rankall <- function(outcome, num = "best") {## Read outcome data
```

```
    hospital_data <- read.csv("outcome-of-care-measures.csv", stringsAsFactors = FALSE)
    colnames(hospital_data)[2] <- "hospital"
    colnames(hospital_data)[7] <- "state"
    colnames(hospital_data)[11] <- "heart attack"
    colnames(hospital_data)[17] <- "heart failure"
    colnames(hospital_data)[23] <- "pneumonia"
    ## Check that state and outcome are valid
    if(!outcome %in% c("heart attack", "heart failure",  "pneumonia")) stop("invalid outcome")
    ##Creating a data frame without NA values
    hospital_data_na <- if(outcome == "heart attack"){
        hospital_data[!is.na(as.numeric(hospital_data$"heart attack")) , ]
    } else if (outcome == "heart failure"){
        hospital_data[!is.na(as.numeric(hospital_data$"heart failure")) , ]
    } else if (outcome == "pneumonia"){
        hospital_data[!is.na(as.numeric(hospital_data$"pneumonia")) , ]
    }
    ##Splitting the data into one data frame per State
    split_data <- split.data.frame(hospital_data_na,hospital_data$state)
    ## For each state, find the hospital of the given rank
    all_states_rows <- if(outcome == "heart attack"){
        lapply(split_data, function(x) x[order(x$"heart attack",x$hospital), ])
    } else if (outcome == "heart failure"){
        lapply(split_data, function(x) x[order(x$"heart failure",x$hospital), ])
    } else if (outcome == "pneumonia"){
```

```
        lapply(split_data, function(x) x[order(x$pneumonia,x$hospital), ])
    }
##NOTE: I'm almost sure that the code above works as expected
    ## Return a data frame with the hospital names and the
    ## (abbreviated) state name
    ranked_hospital <- if(num == "best"){
        hospital <- lapply(all_states_rows, function(x) x[1,2])
        state <- lapply(all_states_rows, function(x) x[1,7])
        as.data.frame(cbind(hospital,state))
    }
    else if (is.numeric(num)){
        hospital <- lapply(all_states_rows, function(x) x[num,2])
        state <- lapply(all_states_rows, function(x) x[num,7])
        as.data.frame(cbind(hospital,state))
    }
    else if (num == "worst"){
        hospital <- lapply(all_states_rows, function(x) x[length(x$hospital),2])
        state <- lapply(all_states_rows, function(x) x[length(x$hospital),7])
        as.data.frame(cbind(hospital,state))
    }
}
```

I can't figure out how to extract each hospital name and state and form the expected data frame as output, so far I've been trying with the apply functions but I can't figure out how to use a loop with my current list ("all_states_rows") as an alternative. I'll appreciate any help. Thanks for your time!!

0 · flag

Anonymous · 9 days ago

Woah, there are a lot of different lapply and conditional statements in there.

In general one approach you can do is this:

1. Split the data, so you have a list of data frames, one for each state
2. Observe that the "names" of that list will be a list of states; i.e `names(splitResult)` will be `c("AL", "AK", "AZ"...)`
3. Call `lapply` on this list with a single function that gets you back the hospital name for each data frame
4. Package the results into the data frame format required

So for #3, in pseudocode...

```
lapplyResult <- lapply(splitData,
    function(e) {
        # anonymous function; e = a single data frame
        # remove NAs...
        # reorder the data frame as required...
        # return either the best hospital, worst hospital, or num-th hospital, depending on "n
```

```
um" argument...
        # (that part is just subsetting; for example, head(e$Hospital.Name, 1) for the best ho
spital...
    })

  # now lapplyResult$AL = (hospitalName or NA), lapplyResult$AK = (hospitalName or NA), ...
  # so package the data frame as the function spec requires.
```

0 · flag

+ Comment

New post

To ensure a positive and productive discussion, please read our forum posting policies before posting.

B | I | Link | <code> | Pic | Math | Edit: Rich | Preview

Make this post anonymous to other students

Subscribe to this thread at the same time

Add post