Programming Assignment 3 - RankAll

Shafeeq Ur Rahaman

3/14/2018

Ranking hospitals in all states

```
rankall <- function(outcome, num="best")
{
    #Read the dataset
    outcome_dataset <- read.csv("outcome-of-care-measures.csv", colClasses = "character")
    #Read the subset of data required
    outcome_subset <- subset(outcome_dataset, select = c(2,7,11,17,23))
    #Assign column names to the dataset
    colnames(outcome_subset) <- c("hospital_name", "state", "heart_attack","heart_failure", "pneumonia"
    #Convert the space between words of outcome for better coding
    outcome <- gsub(' ','_',outcome)
    #Check for outcome validity
    if(!(outcome %in% colnames(outcome_subset[,3:5])))
    {
        stop("Invalid Outcome")
    }
    else if(is.numeric(num)) #Check if num is numeric
    {
        outcome_subset[,outcome] <- as.numeric(outcome_subset[,outcome])
        #Remove the NA values of outcome
        outcome_subset <- outcome_subset[complete.cases(outcome_subset[,outcome]),]
        #Create list of dataframes (states)
        outcome_split <- split(outcome_subset, outcome_subset$state)
        #outcome_rank_filter <- list() #To hold the filtered ranks
        output <- do.call(rbind, lapply(outcome_split, function(x)
            {
         x <- x[order(x$hospital_name), ]
         x <- x %>% mutate(rank = rank(x[, outcome], ties.method = "first"))
         x %>% filter(rank == num) %>% select(hospital_name, state)
            })
          )
        #Find the unique states so as to replace the missing states with NA
        state_names <- unique(outcome_subset$state)
        #Replace the missing states with NA and update the output
        output <- output %>% mutate(state = factor(state, levels = state_names)) %>% complete(state, fi
        #Convert tibble to data frame in order to provide row names
        output <- as.data.frame(lapply(output,as.character), stringsAsFactors = FALSE)
        #sort the state
        output <- output[order(output$state),]
        #Rearrange the variables to meet the requirements
        output <- output[,c(2,1)]
        #Add rownames as state to the data frame
        rownames(output) <- output[,2]
    }
    else if(!is.numeric(num))
```

```
{
    if(num=="best")
    {

        outcome_subset[,outcome] <- as.numeric(outcome_subset[,outcome])
        #Remove the NA values of outcome
        outcome_subset <- outcome_subset[complete.cases(outcome_subset[,outcome]),]
        #Create list of dataframes (states)
        outcome_split <- split(outcome_subset, outcome_subset$state)
        outcome_rank_filter <- list() #To hold the filtered ranks
        for( i in seq_along(outcome_split))
        {
            #Ordered the data frames in the list by hospital name
            outcome_split[[i]] <- outcome_split[[i]][order(outcome_split[[i]]$hospital_name),]
            #Add Rank column for each data frame in the list
            outcome_split[[i]] <- outcome_split[[i]] %>% mutate(rank = rank(outcome_split[[i]][,out
            #Filter the dataset based on the rank (num) passed to the function
            outcome_rank_filter[[i]] <- outcome_split[[i]] %>% arrange(rank) %>% filter(rank == min
        }
        #Collapse the list of data frame into one data frame
        output <- do.call(rbind, outcome_rank_filter)
        #Add rownames as state to the data frame
        rownames(output) <- output[,2]
    }
    else if(num=="worst")
    {
        outcome_subset[,outcome] <- as.numeric(outcome_subset[,outcome])
        #Remove the NA values of outcome
        outcome_subset <- outcome_subset[complete.cases(outcome_subset[,outcome]),]
        #Create list of dataframes (states)
        outcome_split <- split(outcome_subset, outcome_subset$state)
        outcome_rank_filter <- list() #To hold the filtered ranks
        for( i in seq_along(outcome_split))
        {
            #Ordered the data frames in the list by hospital name
            outcome_split[[i]] <- outcome_split[[i]][order(outcome_split[[i]]$hospital_name),]
            #Add Rank column for each data frame in the list
            outcome_split[[i]] <- outcome_split[[i]] %>% mutate(rank = rank(outcome_split[[i]][,out
            #Filter the dataset based on the rank (num) passed to the function
            outcome_rank_filter[[i]] <- outcome_split[[i]] %>% arrange(rank) %>% filter(rank == max
        }
        #Collapse the list of data frame into one data frame
        output <- do.call(rbind, outcome_rank_filter)
        #Add rownames as state to the data frame
        rownames(output) <- output[,2]
    }
}
else
{
    stop("Invalid num")
}
return(output)
```

```
}
```

```
head(rankall("heart attack", 20), 10)
```

```
##                           hospital_name state
## AK                                 <NA>    AK
## AL      D W MCMILLAN MEMORIAL HOSPITAL    AL
## AR   ARKANSAS METHODIST MEDICAL CENTER    AR
## AZ JOHN C LINCOLN DEER VALLEY HOSPITAL    AZ
## CA               SHERMAN OAKS HOSPITAL    CA
## CO            SKY RIDGE MEDICAL CENTER    CO
## CT             MIDSTATE MEDICAL CENTER    CT
## DC                                 <NA>    DC
## DE                                 <NA>    DE
## FL      SOUTH FLORIDA BAPTIST HOSPITAL    FL
```

```
tail(rankall("pneumonia", "worst"), 3)
```

```
##                                 hospital_name state
## WI MAYO CLINIC HEALTH SYSTEM - NORTHLAND, INC    WI
## WV                     PLATEAU MEDICAL CENTER    WV
## WY           NORTH BIG HORN HOSPITAL DISTRICT    WY
```

```
tail(rankall("heart failure"), 10)
```

```
##                                                    hospital_name state
## TN                          WELLMONT HAWKINS COUNTY MEMORIAL HOSPITAL    TN
## TX                                       FORT DUNCAN MEDICAL CENTER    TX
## UT VA SALT LAKE CITY HEALTHCARE - GEORGE E. WAHLEN VA MEDICAL CENTER    UT
## VA                                        SENTARA POTOMAC HOSPITAL    VA
## VI                              GOV JUAN F LUIS HOSPITAL & MEDICAL CTR    VI
## VT                                          SPRINGFIELD HOSPITAL    VT
## WA                                       HARBORVIEW MEDICAL CENTER    WA
## WI                                  AURORA ST LUKES MEDICAL CENTER    WI
## WV                                       FAIRMONT GENERAL HOSPITAL    WV
## WY                                      CHEYENNE VA MEDICAL CENTER    WY
```

```
head(rankall("pneumonia", "best"), 3)
```

```
##                           hospital_name state
## AK YUKON KUSKOKWIM DELTA REG HOSPITAL    AK
## AL      MARSHALL MEDICAL CENTER NORTH    AL
## AR        STONE COUNTY MEDICAL CENTER    AR
```

```
r <- rankall("heart attack", 4)
as.character(subset(r, state == "HI")$hospital)
```

```
## [1] "CASTLE MEDICAL CENTER"
```

```
r <- rankall("pneumonia", "worst")
as.character(subset(r, state == "NJ")$hospital)
```

```
## [1] "BERGEN REGIONAL MEDICAL CENTER"
```

```
r <- rankall("heart failure", 10)
as.character(subset(r, state == "NV")$hospital)
```

```
## [1] "RENOWN SOUTH MEADOWS MEDICAL CENTER"
```