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$state)  
 #Replace the missing states with NA and update the output  
 output <- output %>% mutate(state = factor(state, levels = state\_names)) %>% complete(state, fill = list(hospital\_name = NA))  
 #Convert tibble to data frame in order to provide row names  
 output <- as.data.frame(lapply(output,as.character), stringsAsFactors = FALSE)  
 #sort the state  
 output <- output[order(output$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]][,outcome], ties.method="first"))   
 #Filter the dataset based on the rank (num) passed to the function  
 outcome\_rank\_filter[[i]] <- outcome\_split[[i]] %>% arrange(rank) %>% filter(rank == min(rank)) %>% select(hospital\_name,state)  
 }  
 #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]][,outcome], ties.method="first"))   
 #Filter the dataset based on the rank (num) passed to the function  
 outcome\_rank\_filter[[i]] <- outcome\_split[[i]] %>% arrange(rank) %>% filter(rank == max(rank)) %>% select(hospital\_name,state)  
 }  
 #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"