Programming Assignment Air quality

Hariharan

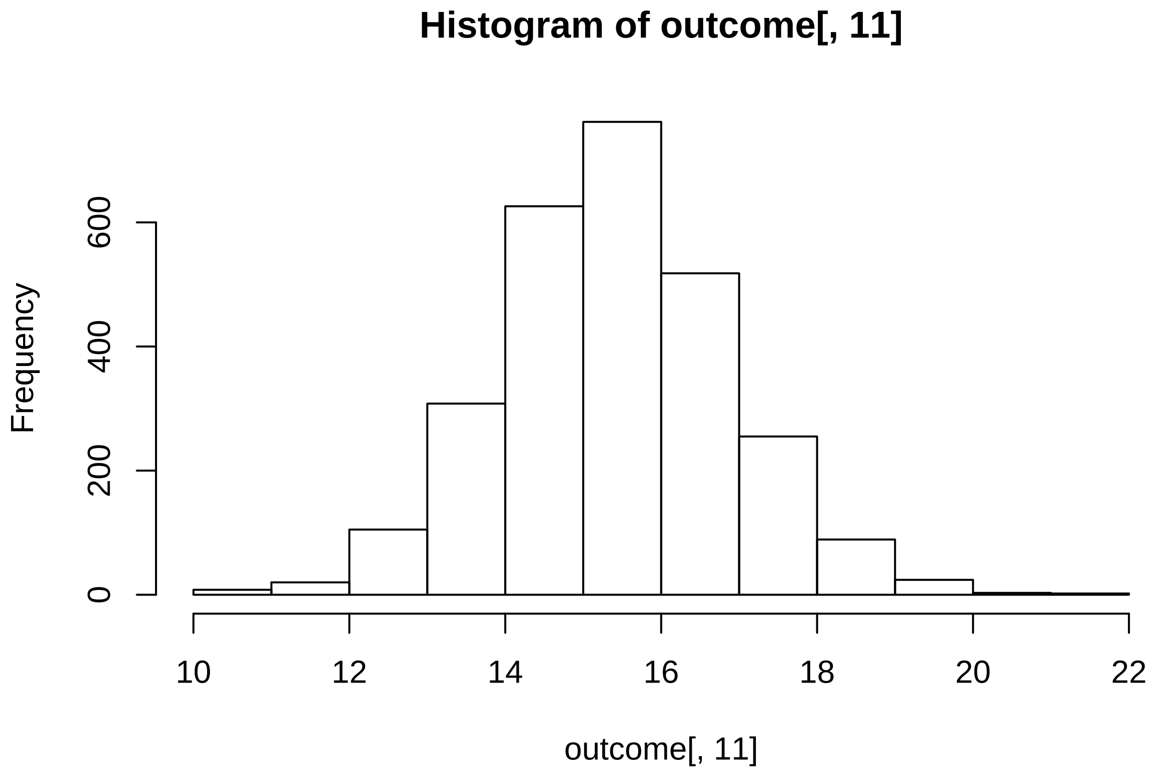
This PDF contains the code and output generated for programming assignment 3 in the course Find the week_2_prog_assignment.Rmd in the same folder as this file to interact with the code and make changes for a better learning experience Download the file ProgAssignment3-data.zip file containing the data for Programming Assignment 3 from the Coursera web site. Unzip the file in a directory that will serve as your working directory. When you start up R make sure to change your working directory to the directory where you unzipped the data.

1 Plot the 30-day mortality rates for heart attack

```
outcome <- read.csv("outcome-of-care-measures.csv", colClasses = "character")
outcome[, 11] <- as.numeric(outcome[, 11])
```

```
## Warning: NAs introduced by coercion
```

```
hist(outcome[, 11])
```



2 Finding the best hospital in a state

Write a function called best that take two arguments: the 2-character abbreviated name of a state and an outcome name. The function reads the outcome-of-care-measures.csv file and returns a character vector with the name of the hospital that has the best (i.e. lowest) 30-day mortality for the specified outcome in that state. The hospital name is the name provided in the Hospital.Name variable. The outcomes can be one of “heart attack”, “heart failure”, or “pneumonia”. Hospitals that do not have data on a particular outcome should be excluded from the set of hospitals when deciding the rankings.

```r
best <- function(state, outcome) {
  ## Read outcome data
  ## Check that state and outcome are valid
  ## Return hospital name in that state with lowest 30-day death
  ## rate
  data <- read.csv("outcome-of-care-measures.csv", colClasses = "character")
  fd   <- as.data.frame(cbind(data[, 2],   # hospital
                              data[, 7],   # state
                              data[, 11],  # heart attack
                              data[, 17],  # heart failure
                              data[, 23]), # pneumonia
                        stringsAsFactors = FALSE)
  colnames(fd) <- c("hospital", "state", "heart attack", "heart failure", "pneumonia")

  if(!state %in% fd[, "state"])
  {
      stop('invalid state')
  }
  else if(!outcome %in% c("heart attack", "heart failure", "pneumonia"))
  {
    stop('invalid outcome')
  }
  else
  {
    si <- which(fd[, "state"] == state)
    ts <- fd[si, ]    # extracting data for the called state
    oi <- as.numeric(ts[,outcome])
    min_val <- min(oi, na.rm = TRUE)
    result  <- ts[, "hospital"][which(oi == min_val)]
    output  <- result[order(result)]
  }
  return(output)

}
```

Sample usage

```
best("SC", "heart attack")
```

```
## Warning in best("SC", "heart attack"): NAs introduced by coercion
```

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

3 Ranking hospitals by outcome in a state

Write a function called rankhospital that takes three arguments: the 2-character abbreviated name of a state (state), an outcome (outcome), and the ranking of a hospital in that state for that outcome (num). The function reads the outcome-of-care-measures.csv file and returns a character vector with the name of the hospital that has the ranking specified by the num argument. For example, the call rankhospital("MD", "heart failure", 5) would return a character vector containing the name of the hospital with the 5th lowest 30-day death rate for heart failure. The num argument can take values "best", "worst", or an integer indicating the ranking (smaller numbers are better). If the number given by num is larger than the number of hospitals in that state, then the function should return NA. Hospitals that do not have data on a particular outcome should be excluded from the set of hospitals when deciding the rankings.

```
rankhospital <- function(state, outcome, rank = "best"){
  ## Read outcome data
  data <- read.csv("outcome-of-care-measures.csv", colClasses = "character")
  fd   <- as.data.frame(cbind(data[, 2],  # hospital
                              data[, 7],  # state
                              data[, 11],  # heart attack
                              data[, 17],  # heart failure
                              data[, 23]), # pneumonia
                        stringsAsFactors = FALSE)
  colnames(fd) <- c("hospital", "state", "heart attack", "heart failure", "pneumonia")

  ## Check that state and outcome are valid
  if (!state %in% fd[, "state"]) {
    stop('invalid state')
  } else if (!outcome %in% c("heart attack", "heart failure", "pneumonia")){
    stop('invalid outcome')
  } else if (is.numeric(rank)) {
    si <- which(fd[, "state"] == state)
    ts <- fd[si, ]                       # extracting dataframe for the called state
    ts[, eval(outcome)] <- as.numeric(ts[, eval(outcome)])
    ts <- ts[order(ts[, eval(outcome)], ts[, "hospital"]), ]
    output <- ts[, "hospital"][rank]
  } else if (!is.numeric(rank)){
    if (rank == "best") {
      output <- best(state, outcome)
    } else if (rank == "worst") {
      si <- which(fd[, "state"] == state)
      ts <- fd[si, ]
      ts[, eval(outcome)] <- as.numeric(ts[, eval(outcome)])
      ts <- ts[order(ts[, eval(outcome)], ts[, "hospital"], decreasing = TRUE), ]
      output <- ts[, "hospital"][1]
    } else {
      stop('invalid rank')
    }
  }
  return(output)
}
```

Sample Usage

```
print(rankhospital("NC", "heart attack", "worst"))
```

```
## Warning in rankhospital("NC", "heart attack", "worst"): NAs introduced by
## coercion
```

```
## [1] "WAYNE MEMORIAL HOSPITAL"
```

4 Ranking hospitals in all states

Write a function called rankall that takes two arguments: an outcome name (outcome) and a hospital ranking (num). The function reads the outcome-of-care-measures.csv file and returns a 2-column data frame containing the hospital in each state that has the ranking specified in num. For example the function call rankall(“heart attack”, “best”) would return a data frame containing the names of the hospitals that are the best in their respective states for 30-day heart attack death rates. The function should return a value for every state (some may be NA). The first column in the data frame is named hospital, which contains the hospital name, and the second column is named state, which contains the 2-character abbreviation for the state name. Hospitals that do not have data on a particular outcome should be excluded from the set of hospitals when deciding the rankings.

```
rankall <- function(outcome, num = "best"){
  ## Read outcome data
  data <- read.csv("outcome-of-care-measures.csv", colClasses = "character")
  fd   <- as.data.frame(cbind(data[, 2],  # hospital
                              data[, 7],  # state
                              data[, 11],  # heart attack
                              data[, 17],  # heart failure
                              data[, 23]), # pneumonia
                        stringsAsFactors = FALSE)
  colnames(fd) <- c("hospital", "state", "heart attack", "heart failure", "pneumonia")
  fd[, eval(outcome)] <- as.numeric(fd[, eval(outcome)])

  ## Check that state and outcome are valid

  if (!outcome %in% c("heart attack", "heart failure", "pneumonia")){
    stop('invalid outcome')
  } else if (is.numeric(num)) {
    by_state <- with(fd, split(fd, state))
    ordered  <- list()
    for (i in seq_along(by_state)){
      by_state[[i]] <- by_state[[i]][order(by_state[[i]][, eval(outcome)],
                                           by_state[[i]][, "hospital"]), ]
      ordered[[i]]  <- c(by_state[[i]][num, "hospital"], by_state[[i]][, "state"][1])
    }
    result <- do.call(rbind, ordered)
    output <- as.data.frame(result, row.names = result[, 2], stringsAsFactors = FALSE)
    names(output) <- c("hospital", "state")
  } else if (!is.numeric(num)) {
    if (num == "best") {
      by_state <- with(fd, split(fd, state))
      ordered  <- list()
      for (i in seq_along(by_state)){
        by_state[[i]] <- by_state[[i]][order(by_state[[i]][, eval(outcome)],
                                             by_state[[i]][, "hospital"]), ]
        ordered[[i]]  <- c(by_state[[i]][1, c("hospital", "state")])
      }
      result <- do.call(rbind, ordered)
      output <- as.data.frame(result, stringsAsFactors = FALSE)
      rownames(output) <- output[, 2]
    } else if (num == "worst") {
      by_state <- with(fd, split(fd, state))
      ordered  <- list()
      for (i in seq_along(by_state)){
        by_state[[i]] <- by_state[[i]][order(by_state[[i]][, eval(outcome)],
                                             by_state[[i]][, "hospital"],
                                             decreasing = TRUE), ]
        ordered[[i]]  <- c(by_state[[i]][1, c("hospital", "state")])
      }
      result <- do.call(rbind, ordered)
      output <- as.data.frame(result, stringsAsFactors = FALSE)
      rownames(output) <- output[, 2]
    } else {
      stop('invalid num')
    }
  }
  return(output)
}
```

Sample Useage

```
print(rankall("heart attack", "worst"))
```

```
## Warning in rankall("heart attack", "worst"): NAs introduced by coercion
```

```
##                                                                 hospital state
## AK                                      MAT-SU REGIONAL MEDICAL CENTER    AK
## AL                                      HELEN KELLER MEMORIAL HOSPITAL    AL
## AR                                       MEDICAL CENTER SOUTH ARKANSAS    AR
## AZ                                         VERDE VALLEY MEDICAL CENTER    AZ
## CA                                    METHODIST HOSPITAL OF SACRAMENTO    CA
## CO                                       NORTH SUBURBAN MEDICAL CENTER    CO
## CT                                           JOHNSON MEMORIAL HOSPITAL    CT
## DC                                          HOWARD UNIVERSITY HOSPITAL    DC
## DE                                               ST FRANCIS HEALTHCARE    DE
## FL                                           PALMETTO GENERAL HOSPITAL    FL
## GA                                         WEST GEORGIA MEDICAL CENTER    GA
## GU                                    GUAM MEMORIAL HOSPITAL AUTHORITY    GU
## HI                                            PALI MOMI MEDICAL CENTER    HI
## IA                                               BOONE COUNTY HOSPITAL    IA
## ID                               EASTERN IDAHO REGIONAL MEDICAL CENTER    ID
## IL                                        SAINT ANTHONY MEDICAL CENTER    IL
## IN                                             MARION GENERAL HOSPITAL    IN
## KS                                               OLATHE MEDICAL CENTER    KS
## KY                                     MURRAY-CALLOWAY COUNTY HOSPITAL    KY
## LA                                             RIVER PARISHES HOSPITAL    LA
## MA                                                      NOBLE HOSPITAL    MA
## MD                                           HARFORD MEMORIAL HOSPITAL    MD
## ME                                           PENOBSCOT VALLEY HOSPITAL    ME
## MI                                               HURLEY MEDICAL CENTER    MI
## MN                                       HEALTHEAST ST JOHN'S HOSPITAL    MN
## MO                                POPLAR BLUFF REGIONAL MEDICAL CENTER    MO
## MS                                SOUTHWEST MS REGIONAL MEDICAL CENTER    MS
## MT                                          BOZEMAN DEACONESS HOSPITAL    MT
## NC                                             WAYNE MEMORIAL HOSPITAL    NC
## ND                                                      ALTRU HOSPITAL    ND
## NE OMAHA VA MEDICAL CENTER (VA NEBRASKA WESTERN IOWA HEALTHCARE SYSTEM)    NE
## NH                                          FRANKLIN REGIONAL HOSPITAL    NH
## NJ                   ROBERT WOOD JOHNSON UNIVERSITY HOSPITAL AT RAHWAY    NJ
## NM                               MOUNTAIN VIEW REGIONAL MEDICAL CENTER    NM
## NV                                             DESERT SPRINGS HOSPITAL    NV
## NY                                               F F THOMPSON HOSPITAL    NY
## OH                              MERCY FRANCISCAN HOSPITAL WESTERN HILLS    OH
## OK                                        MERCY MEMORIAL HEALTH CENTER    OK
## OR                                     THREE RIVERS COMMUNITY HOSPITAL    OR
## PA                                           EPHRATA COMMUNITY HOSPITAL    PA
## PR                                        DOCTORS' CENTER HOSPITAL, INC    PR
## RI                                                    WESTERLY HOSPITAL    RI
## SC                                          WACCAMAW COMMUNITY HOSPITAL    SC
## SD                                               PRAIRIE LAKES HOSPITAL    SD
## TN                                    DYERSBURG REGIONAL MEDICAL CENTER    TN
## TX                                                LAREDO MEDICAL CENTER    TX
## UT                                                    ST MARKS HOSPITAL    UT
## VA                                      RIVERSIDE TAPPAHANNOCK HOSPITAL    VA
## VI                               GOV JUAN F LUIS HOSPITAL & MEDICAL CTR    VI
## VT                               NORTHEASTERN VERMONT REGIONAL HOSPITAL    VT
## WA                                       KADLEC REGIONAL MEDICAL CENTER    WA
## WI                                             HOLY FAMILY MEMORIAL INC    WI
## WV                                             THOMAS MEMORIAL HOSPITAL    WV
## WY                                           SHERIDAN MEMORIAL HOSPITAL    WY
```

Quiz

```r
best("SC", "heart attack")
```

```
## Warning in best("SC", "heart attack"): NAs introduced by coercion
```

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

```r
best("NY", "pneumonia")
```

```
## Warning in best("NY", "pneumonia"): NAs introduced by coercion
```

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

```r
best("AK", "pneumonia")
```

```
## Warning in best("AK", "pneumonia"): NAs introduced by coercion
```

```
## [1] "YUKON KUSKOKWIM DELTA REG HOSPITAL"
```

```r
rankhospital("NC", "heart attack", "worst")
```

```
## Warning in rankhospital("NC", "heart attack", "worst"): NAs introduced by
## coercion
```

```
## [1] "WAYNE MEMORIAL HOSPITAL"
```

```r
rankhospital("WA", "heart attack", 7)
```

```
## Warning in rankhospital("WA", "heart attack", 7): NAs introduced by coercion
```

```
## [1] "YAKIMA VALLEY MEMORIAL HOSPITAL"
```

```r
rankhospital("TX", "pneumonia", 10)
```

```
## Warning in rankhospital("TX", "pneumonia", 10): NAs introduced by coercion

## [1] "SETON SMITHVILLE REGIONAL HOSPITAL"
```

```r
rankhospital("NY", "heart attack", 7)
```

```
## Warning in rankhospital("NY", "heart attack", 7): NAs introduced by coercion

## [1] "BELLEVUE HOSPITAL CENTER"
```

```r
r <- rankall("heart attack", 4)
```

```
## Warning in rankall("heart attack", 4): NAs introduced by coercion
```

```r
as.character(subset(r, state == "HI")$hospital)
```

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

```r
r <- rankall("pneumonia", "worst")
```

```
## Warning in rankall("pneumonia", "worst"): NAs introduced by coercion
```

```r
as.character(subset(r, state == "NJ")$hospital)
```

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

```r
r <- rankall("heart failure", 10)
```

```
## Warning in rankall("heart failure", 10): NAs introduced by coercion
```

```r
as.character(subset(r, state == "NV")$hospital)
```

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