Quiz 1 Week 1 R Markdown file

Me

4/5/2020

## Reading in the data

### Reading in the data method 1

getwd()

## [1] "/Users/sadikafhameed/Documents/Coursera/R Programming /Week 1/Quiz 1"

data <- read.csv("hw1\_data.csv")

### Reading in the data method 2

dat <- download.file('https://d396qusza40orc.cloudfront.net/rprog/data/quiz1\_data.zip', destfile ="quizdat.zip")  
dat <- unzip("quizdat.zip")  
dat <- read.csv("hw1\_data.csv")

## In the dataset provided for this Quiz, what are the column names of the dataset?

### MEethod 1

colnames(data)

## [1] "Ozone" "Solar.R" "Wind" "Temp" "Month" "Day"

### Method 2

head(data)

## Ozone Solar.R Wind Temp Month Day  
## 1 41 190 7.4 67 5 1  
## 2 36 118 8.0 72 5 2  
## 3 12 149 12.6 74 5 3  
## 4 18 313 11.5 62 5 4  
## 5 NA NA 14.3 56 5 5  
## 6 28 NA 14.9 66 5 6

### Method 3

names(dat)

## [1] "Ozone" "Solar.R" "Wind" "Temp" "Month" "Day"

## Extract the first 2 rows of the data frame and print them to the console. What does the output look like?

### Method 1

dat[1:2, ]

## Ozone Solar.R Wind Temp Month Day  
## 1 41 190 7.4 67 5 1  
## 2 36 118 8.0 72 5 2

### Method 2

dat[152:153, ]

## Ozone Solar.R Wind Temp Month Day  
## 152 18 131 8.0 76 9 29  
## 153 20 223 11.5 68 9 30

### Method 3

tail(data, 2)

## Ozone Solar.R Wind Temp Month Day  
## 152 18 131 8.0 76 9 29  
## 153 20 223 11.5 68 9 30

## How many observations (i.e. rows) are in this data frame?

nrow(data)

## [1] 153

## Question 15 What is the value of Ozone in the 47th row?

### Method 1

``` data[47, ]

### 

dat$Ozone[47]

## [1] 21

## How many missing values are in the Ozone column of this data frame?

### Method 1

sum(is.na(dat$Ozone))

## [1] 37

### Method 2

sub <- subset(data, is.na(Ozone))  
nrow(sub)

## [1] 37

### Method 3

#### To calculate as an example how many missing values there are in total

length(which(is.na((data))))

## [1] 44

#### To calculate only for Ozone

datana <- subset(dat, is.na(Ozone))  
nrow(datana)  
  
## What is the mean of the Ozone column in this dataset? Exclude missing values (coded as NA) from this calculation.  
  
### Method 1

datanotna <- subset(dat, !is.na(Ozone)) apply(datanotna, 2, mean)

### Method 2

``` mean(dat$Ozone, na.rm = TRUE) ### Method 3

sub = subset(data, !is.na(Ozone), select = Ozone)  
apply(sub, 2, mean)

## Ozone   
## 42.12931

## Extract the subset of rows of the data frame where Ozone values are above 31 and Temp values are above 90. What is the mean of Solar.R in this subset?

### Method 1

sub <- subset(data, Ozone > 31 & Temp >90, select = Solar.R)  
apply(sub, 2, mean)

## Solar.R   
## 212.8

### Method 2

mean(dat[which(dat$Ozone >31 & dat$Temp > 90),]$Solar.R)

## [1] 212.8

### Method 3

datasub <- subset(dat, dat$Ozone > 31 & dat$Temp >90, select = Solar.R)  
apply(datasub, 2, mean)

## Solar.R   
## 212.8

## What is the mean of “Temp” when “Month” is equal to 6?

### Method 1

mean(dat[which(dat$Month == 6), ]$Temp)

### Method 2

sub = subset(dat, Month == 6, select = Temp)  
apply(sub, 2, mean)

### Method 3

datasub <- subset(data, data$Month == 6, select = Temp)  
apply(datasub, 2 , mean)

## What was the maximum ozone value in the month of May (i.e. Month = 5)

### Method 1

datasub <- subset(data, !is.na(Ozone) & data$Month ==5, select = Ozone)  
apply(datasub, 2, max)

### Method 2

sub <- subset(data, Month == 5 & !is.na(Ozone), select = Ozone )  
apply(sub, 2, max)