Project 4 Solutions

(Abhimanyu Agarwal)

Collaborators: N/A

TA help:

1) Hilda Ibriga : Helped me with Question 5

Online resources used: N/A

Question 1

```
### Question 1
#Loads into dataframe called "splash_mountain" using read.csv()
splash_mountain <- read.csv("/class/datamine/data/disney/splash_mountain.csv")

#Mean using loop along with conditionals
counter <- 0                                    #Counter variable
sum <- 0
for(i in splash_mountain$SPOSTMIN)              #Enters the for loop, iterates over SPOSTMIN column
  {
  if(is.na(i) != TRUE )                         #Makes sure NA values arent included in the mean sum
    {
    if( i != -999)                              #Makes sure -999 values arent included in the mean sum
      {
      sum <- sum + i                            #Add values one by one to the sum
      counter <- counter + 1                    #Counter variable increments by 1
      }
    }
  }
mean <- sum/counter                             #Computes the mean as shown by the division
head(mean)                                      #Prints the mean value
```

```
[1] 43.3892
```

Question 2

```
#Loads into dataframe called "flight_of_passage" using read.csv()
ride_name <- read.csv("/class/datamine/data/disney/flight_of_passage.csv")
for (i in 1:nrow(ride_name))                                          #Enters the for loop, iterates
  {
  if(any(ride_name[i, c("SPOSTMIN", "SACTMIN")] == -999, na.rm = TRUE)) #Checks conditions to classify
    {
      ride_name$status[i] = "closed"                                  #Assign the status "closed" if
    }
  else
    {
      ride_name$status[i] = "open"                                    #Assign the status "open" if va
```

```
    }
  }
ride_name$status = factor(ride_name$status)                              #After exiting the loop, it cha

#Printing str result
str(ride_name)
```

```
'data.frame':   122163 obs. of  5 variables:
 $ date    : chr  "05/26/2017" "05/26/2017" "05/26/2017" "05/26/2017" ...
 $ datetime: chr  "2017-05-26 09:06:38" "2017-05-26 09:10:11" "2017-05-26 09:10:12" "2017-05-26 09:17:09
 $ SACTMIN : int  NA 47 NA NA NA NA NA NA NA NA ...
 $ SPOSTMIN: int  -999 NA 5 60 60 60 45 45 45 45 ...
 $ status  : Factor w/ 2 levels "closed","open": 1 2 2 2 2 2 2 2 2 2 ...
```

```
table(ride_name$status)                                                  #Prints the number of open and
```

```
closed   open
  4637 117526
```

Question 3

```
#Assigns status open
status <- rep("open", times = nrow(ride_name))

#Applies conditionals and assigns "closed" status
status[which((ride_name$SPOSTMIN) == -999 | (ride_name$SACTMIN == -999))] <- "closed"

#After checking the conditionals, it changes it into factor
ride_name$status <- factor(status)

#Displays the entire dataframe, you could check the new status
head(ride_name)
```

```
        date            datetime SACTMIN SPOSTMIN status
1 05/26/2017 2017-05-26 09:06:38      NA     -999 closed
2 05/26/2017 2017-05-26 09:10:11      47       NA   open
3 05/26/2017 2017-05-26 09:10:12      NA        5   open
4 05/26/2017 2017-05-26 09:17:09      NA       60   open
5 05/26/2017 2017-05-26 09:24:07      NA       60   open
6 05/26/2017 2017-05-26 09:30:10      NA       60   open
```

```
#Printing str result
str(ride_name)
```

```
'data.frame':   122163 obs. of  5 variables:
 $ date    : chr  "05/26/2017" "05/26/2017" "05/26/2017" "05/26/2017" ...
 $ datetime: chr  "2017-05-26 09:06:38" "2017-05-26 09:10:11" "2017-05-26 09:10:12" "2017-05-26 09:17:09
 $ SACTMIN : int  NA 47 NA NA NA NA NA NA NA NA ...
 $ SPOSTMIN: int  -999 NA 5 60 60 60 45 45 45 45 ...
 $ status  : Factor w/ 2 levels "closed","open": 1 2 2 2 2 2 2 2 2 2 ...
```

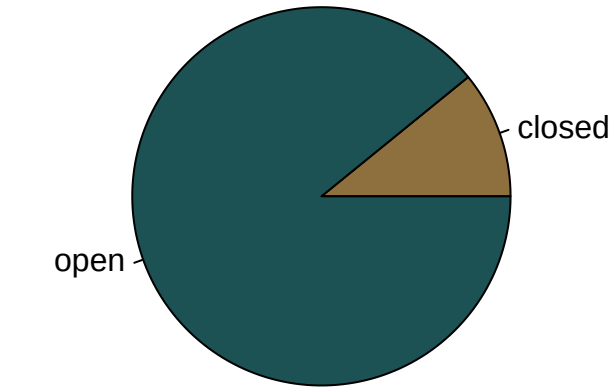
```
table(ride_name$status)
```

```
closed   open
```

```
  4637 117526
```

Question 4

```
#Using table to get a count of "closed" and "open" status
#Loads into dataframe called "splash_mountain" using read.csv()
splash_mountain <- read.csv("/class/datamine/data/disney/splash_mountain.csv")

#Assigns status open
status <- rep("open", times = nrow(splash_mountain))

#Applies conditionals and assigns "closed" status
status[which((splash_mountain$SPOSTMIN) == -999 | (splash_mountain$SACTMIN == -999))] <- "closed"

#After checking the conditionals, it changes it into factor
splash_mountain$status <- factor(status)

pie(table(splash_mountain$status), col = c("#8E6F3E", "#1c5253"), main = "Updated Status for Splash Mou
```

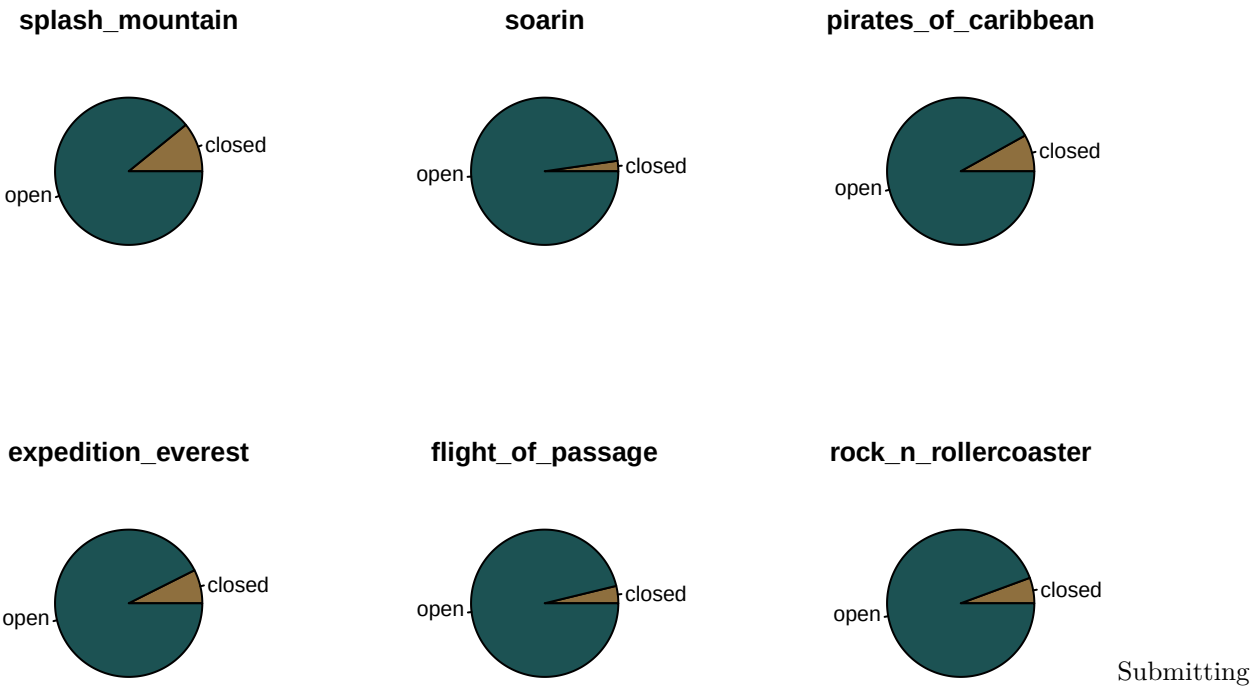



```
#Obtain closed vs open categorical values
#plot the data into pie chart
```

###Question 5

```
ride_names <- c("splash_mountain", "soarin", "pirates_of_caribbean", "expedition_everest", "flight_of_pa
ride_files <- paste0(c("/class/datamine/data/disney/"), ride_names, ".csv")

mypiechart <- function(x){

#Saving it into a "ride file" dataframe
  ride_file <- read.csv(paste0(c("/class/datamine/data/disney/"), x, ".csv"))

#Assigns status open
  ride_file$status <- "open"

#Applies conditionals and assigns "closed" status
  ride_file$status[which(ride_file$SPOSTMIN == -999 | ride_file$SACTMIN == -999)] <- "closed"
```

```
#After checking the conditionals, it changes it into factor
  ride_file$status = factor(ride_file$status)

#Pie chart and its formatting
  pie(table(ride_file$status), col = c("#8E6F3E", "#1c5253"), main = x)
}

par(mfrow=c(2,3))
for (i in ride_names)
  {
  mypiechart(i)
  }
```



Submitting deliverables: project04.RMD, project04.R and project04.pdf

Pledge

By submitting this work I hereby pledge that this is my own, personal work. I've acknowledged in the designated place at the top of this file all sources that I used to complete said work, including but not limited to: online resources, books, and electronic communications. I've noted all collaboration with fellow students and/or TA's. I did not copy or plagiarize another's work.

> As a Boilermaker pursuing academic excellence, I pledge to be honest and true in all that I do. Accountable together - We are Purdue.