Homework 2

Kevin Jin

2/20/2022

Problem 1

```
# (a)
scores <- matrix(c(34, 23, 53, 6, 78, 93, 12, 41, 99),
                 nrow = 3)
scores <- as.data.frame(scores)
names(scores) <- c("car_score",
                   "van_score",
                   "truck_score")

# (b)
library(ggplot2)
head(mpg)
```

```
## # A tibble: 6 x 11
##   manufacturer model displ  year   cyl trans      drv     cty   hwy fl    class
##   <chr>        <chr> <dbl> <int> <int> <chr>      <chr> <int> <int> <chr> <chr>
## 1 audi         a4      1.8  1999     4 auto(l5)   f        18    29 p     compa~
## 2 audi         a4      1.8  1999     4 manual(m5) f        21    29 p     compa~
## 3 audi         a4      2    2008     4 manual(m6) f        20    31 p     compa~
## 4 audi         a4      2    2008     4 auto(av)   f        21    30 p     compa~
## 5 audi         a4      2.8  1999     6 auto(l5)   f        16    26 p     compa~
## 6 audi         a4      2.8  1999     6 manual(m5) f        18    26 p     compa~
```

```
mpg_new <- mpg[mpg$cyl <- 6, ]
mpg_new$class <- as.character(mpg_new$class)
```

Problem 2

```
# read TA data into data frame
tae <- read.table("tae.data",
                  sep = ",")

# create ta_id variable and assign id to each TA sequentially
for (row in 1:nrow(tae)) {
  tae$ta_id[row] <- row
}

# assign names to each variable
```

```r
names(tae) <- c("eng_speaker",
                "instructor_id",
                "course_id",
                "regular",
                "size",
                "score",
                "ta_id")

# (a) coerce eng_speaker variable to logical
# 1 = English speaker -> (TRUE, 1)
# 2 = Non-English speaker -> (FALSE, 0)
for (row in 1:nrow(tae)) {
  if (tae$eng_speaker[row] == 2) {
    tae$eng_speaker[row] <- 0
  }
}
tae$eng_speaker <- as.logical(tae$eng_speaker)

# (b) coerce regular (semester) variable to logical
# 1 = Summer semester -> (FALSE, 0)
# 2 = Regular semester -> (TRUE, 1)
for (row in 1:nrow(tae)) {
  if (tae$regular[row] == 1) {
    tae$regular[row] <- 0
  }
}
tae$regular <- as.logical(tae$regular)

# (c) coerce score variable to factor
# 1 = low
# 2 = medium
# 3 = high
tae$score <- factor(tae$score,
                    levels = c(1, 2, 3),
                    labels = c("Low", "Medium", "High"))

# (d) mean and median class sizes in regular and summer
# mean and median class size in regular semester
reg <- which(tae$regular == TRUE)
round(mean(tae$size[reg]), 2)
```

```
## [1] 29.34
```

```r
median(tae$size[reg])
```

```
## [1] 29
```

```r
# mean and median class size in summer semester
sum <- which(tae$regular == FALSE)
round(mean(tae$size[sum]), 2)
```

```
## [1] 19.7
```

```r
median(tae$size[sum])
```

```
## [1] 20
```

```r
# (e) number of native English and non-native TAs in regular and summer
eng_reg <- which(tae$eng_speaker == TRUE & tae$regular == TRUE)
non_eng_reg <- which(tae$eng_speaker == FALSE & tae$regular == TRUE)
eng_sum <- which(tae$eng_speaker == TRUE & tae$regular == FALSE)
non_eng_sum <- which(tae$eng_speaker == FALSE & tae$regular == FALSE)

# native TAs in regular semester
length(tae$eng_speaker[eng_reg])
```

```
## [1] 20
```

```r
# non-native TAs in regular semester
length(tae$eng_speaker[non_eng_reg])
```

```
## [1] 108
```

```r
# native TAs in summer semester
length(tae$eng_speaker[eng_sum])
```

```
## [1] 9
```

```r
# non-native TAs in summer semester
length(tae$eng_speaker[non_eng_sum])
```

```
## [1] 14
```

Problem 3

Throughout my college career, I have been fortunate enough to largely avoid hitchhiker or couch potato team members in my group projects; however, I have encountered them before, mostly in lower level undergraduate classes. They have taken advantage of me before, because a few years ago, I fell in the category described in the paper as "unwilling to allow a slacker to fail" and "able to cooperate but not delegate". Since then, I have matured and gained more academic experience, and I have also learned the importance of being firm. This has helped me succeed even when my team members are equally as hardworking as I am, because it increases our efficiency even more. To prevent slacker behavior from affecting my performance in this class, I will help establish firm deadlines with my group, keep constant communication through group chat, set a consistent schedule of meetings, and break up large tasks into smaller ones to be completed between meetings. In case anything goes wrong, I will first attempt to communicate and be firm in resolving our differences, and then if my attempts fail, go to the professor to petition to fire the offending student.

Problem 4

1. Team Name: Go Comets!
2. Team Member Info:

Name	Major	Class standing
Michael Tsang	Actuarial Science	Sophomore
Mingyu Sun	Actuarial Science	Senior
Kevin Jin	Data Science	Non-Degree Student

3. Our first meeting was on February 18, 2022. Our next meeting will be on February 25, 2022.