R Homework Number 7

Fisher Ankney

November 6th, 2018

Statistics 5193

*Note: this document was created using R Markdown.

Question 1.

```
library(readxl)
StudentData <- read_excel("/Users/fisher/Documents/data_science/r_stat_5193/data/StudentData.xlsx")
fra <- data.frame(StudentData)
row.names(fra) <- seq(1:35)
```

Question 2.

```
fra$TxtSent[1] <- NA
fra$TxtRec[1] <- NA

fra[1:3,]
```

```
##   Gender    Class HSClass TxtSent TxtRec Fbtime Pinterest Snapchat
## 1      M STAT2023       1      NA     NA     30         N        Y
## 2      M STAT2023      15      10     15     20         N        Y
## 3      M STAT2023      65     150    150     80         N        Y
##   Introvert
## 1         8
## 2         8
## 3         1
```

Question 3.

```
d.text <- fra$TxtSent - fra$TxtRec
fra <- data.frame(fra[1:5], d.text, fra[6:9])

fra[1:3,]
```

```
##   Gender    Class HSClass TxtSent TxtRec d.text Fbtime Pinterest Snapchat
## 1      M STAT2023       1      NA     NA     NA     30         N        Y
## 2      M STAT2023      15      10     15     -5     20         N        Y
## 3      M STAT2023      65     150    150      0     80         N        Y
##   Introvert
## 1         8
## 2         8
## 3         1
```

Question 4.

```
d.text.cat <- factor(sign(d.text), ordered = TRUE,
                     levels = c(-1, 0, 1),
                     labels = c("Neg", "0", "Pos"))

fra$d.text.cat <- d.text.cat

addNA(fra$d.text.cat)
```

```
fra
```

```
##    Gender    Class HSClass TxtSent TxtRec d.text Fbtime Pinterest Snapchat
## 1       M STAT2023       1      NA     NA     NA     30         N        Y
## 2       M STAT2023      15      10     15     -5     20         N        Y
## 3       M STAT2023      65     150    150      0     80         N        Y
## 4       F STAT2023     123      18     28    -10     45         N        Y
## 5       F STAT2023     130      30     30      0     20         N        Y
## 6       F STAT2023     140     100     75     25     60         Y        Y
## 7       M STAT2023     142      20     20      0      5         N        Y
## 8       F STAT2023     200     100    100      0     20         Y        Y
## 9       F STAT2023     200     150    150      0     20         Y        Y
## 10      F STAT2023     220      75     75      0     20         Y        Y
## 11      F STAT2023     220      75     75      0     60         Y        Y
## 12      M STAT2023     258      50     50      0      5         N        Y
## 13      M STAT2023     326      25     30     -5      5         N        N
## 14      F STAT2023     420     100    100      0     60         Y        Y
## 15      M STAT2023     475      30     30      0     20         N        Y
## 16      F STAT2023     728       5      5      0      0         N        N
## 17      F STAT2023     760      20     40    -20     10         Y        Y
## 18      F STAT2023     776     200    200      0     60         Y        Y
## 19      F STAT2023     800      50     50      0     30         N        Y
## 20      F STAT2023     850      10     10      0    120         Y        Y
## 21      M STAT2023     865     100    100      0     30         N        Y
## 22      F STAT2023    1127     100    100      0     25         Y        Y
## 23      F STAT5063     625      30     30      0    120         Y        Y
```

```
## 24      F STAT5063       5      10      10       0       0         N         N
## 25      F STAT5063     250      30      30       0      50         Y         N
## 26      F STAT5063      40      10      20     -10       5         N         Y
## 27      M STAT5063      25       9       3       6      10         Y         Y
## 28      M STAT5063     417     100     200    -100      30         N         Y
## 29      F STAT5063      50      11      15      -4      80         N         N
## 30      F STAT5063      50       2       3      -1      10         N         N
## 31      M STAT5063     330      10      10       0      60         N         N
## 32      F STAT5063     160       5       5       0      15         N         N
## 33      M STAT5063     206      10      15      -5      30         Y         N
## 34      F STAT5063      50       5       5       0       5         N         N
## 35      F STAT5063     200      25      25       0      40         Y         N
##    Introvert d.text.cat
## 1        8.0       <NA>
## 2        8.0        Neg
## 3        1.0          0
## 4        4.0        Neg
## 5        4.0          0
## 6        6.0        Pos
## 7        5.0          0
## 8        6.0          0
## 9        3.0          0
## 10       5.0          0
## 11       5.0          0
## 12       3.5          0
## 13       8.0        Neg
## 14       5.0          0
## 15       5.0          0
## 16       8.0          0
## 17       3.0        Neg
## 18       3.0          0
## 19       8.0          0
## 20       3.0          0
## 21       1.0          0
## 22       4.0          0
## 23       7.0          0
## 24       5.0          0
## 25       7.0          0
## 26       5.0        Neg
## 27       5.0        Pos
## 28       6.0        Neg
## 29       3.0        Neg
## 30       4.0        Neg
## 31       3.0          0
## 32       2.0          0
## 33       6.0        Neg
## 34       3.0          0
## 35       7.0          0
```

Question 5.

```
Ordered <- fra[order(fra$Gender,fra$d.text),]
Ordered
```

```
##    Gender    Class HSClass TxtSent TxtRec d.text Fbtime Pinterest Snapchat
## 17      F STAT2023     760      20     40    -20     10         Y        Y
## 4       F STAT2023     123      18     28    -10     45         N        Y
## 26      F STAT5063      40      10     20    -10      5         N        Y
## 29      F STAT5063      50      11     15     -4     80         N        N
## 30      F STAT5063      50       2      3     -1     10         N        N
## 5       F STAT2023     130      30     30      0     20         N        Y
## 8       F STAT2023     200     100    100      0     20         Y        Y
## 9       F STAT2023     200     150    150      0     20         Y        Y
## 10      F STAT2023     220      75     75      0     20         Y        Y
## 11      F STAT2023     220      75     75      0     60         Y        Y
## 14      F STAT2023     420     100    100      0     60         Y        Y
## 16      F STAT2023     728       5      5      0      0         N        N
## 18      F STAT2023     776     200    200      0     60         Y        Y
## 19      F STAT2023     800      50     50      0     30         N        Y
## 20      F STAT2023     850      10     10      0    120         Y        Y
## 22      F STAT2023    1127     100    100      0     25         Y        Y
## 23      F STAT5063     625      30     30      0    120         Y        Y
## 24      F STAT5063       5      10     10      0      0         N        N
## 25      F STAT5063     250      30     30      0     50         Y        N
## 32      F STAT5063     160       5      5      0     15         N        N
## 34      F STAT5063      50       5      5      0      5         N        N
## 35      F STAT5063     200      25     25      0     40         Y        N
## 6       F STAT2023     140     100     75     25     60         Y        Y
## 28      M STAT5063     417     100    200   -100     30         N        Y
## 2       M STAT2023      15      10     15     -5     20         N        Y
## 13      M STAT2023     326      25     30     -5      5         N        N
## 33      M STAT5063     206      10     15     -5     30         Y        N
## 3       M STAT2023      65     150    150      0     80         N        Y
## 7       M STAT2023     142      20     20      0      5         N        Y
## 12      M STAT2023     258      50     50      0      5         N        Y
## 15      M STAT2023     475      30     30      0     20         N        Y
## 21      M STAT2023     865     100    100      0     30         N        Y
## 31      M STAT5063     330      10     10      0     60         N        N
## 27      M STAT5063      25       9      3      6     10         Y        Y
## 1       M STAT2023       1      NA     NA     NA     30         N        Y
##    Introvert d.text.cat
## 17       3.0        Neg
## 4        4.0        Neg
## 26       5.0        Neg
## 29       3.0        Neg
## 30       4.0        Neg
## 5        4.0          0
## 8        6.0          0
## 9        3.0          0
## 10       5.0          0
## 11       5.0          0
## 14       5.0          0
## 16       8.0          0
```

```
## 18       3.0          0
## 19       8.0          0
## 20       3.0          0
## 22       4.0          0
## 23       7.0          0
## 24       5.0          0
## 25       7.0          0
## 32       2.0          0
## 34       3.0          0
## 35       7.0          0
## 6        6.0        Pos
## 28       6.0        Neg
## 2        8.0        Neg
## 13       8.0        Neg
## 33       6.0        Neg
## 3        1.0          0
## 7        5.0          0
## 12       3.5          0
## 15       5.0          0
## 21       1.0          0
## 31       3.0          0
## 27       5.0        Pos
## 1        8.0       <NA>
```

Question 6.

```
U.OSU <- rep('Y', 35)
U.OSU[23:24] <- 'N'
U.OSU[26:29] <- 'N'
U.OSU[31:35] <- 'N'
U.OSU <- data.frame(U.OSU)

index <- as.numeric(row.names(Ordered)) # used row.names to index

Ordered <- cbind(Ordered, U.OSU[index,])
Ordered[1:5,]
```

```
##    Gender    Class HSClass TxtSent TxtRec d.text Fbtime Pinterest Snapchat
## 17      F STAT2023     760      20     40    -20     10         Y        Y
## 4       F STAT2023     123      18     28    -10     45         N        Y
## 26      F STAT5063      40      10     20    -10      5         N        Y
## 29      F STAT5063      50      11     15     -4     80         N        N
## 30      F STAT5063      50       2      3     -1     10         N        N
##    Introvert d.text.cat U.OSU[index, ]
## 17         3        Neg              Y
## 4          4        Neg              Y
## 26         5        Neg              N
## 29         3        Neg              N
## 30         4        Neg              Y
```

Question 7.

```
set.seed(1)
mean(fra$HSClass[sample(1:nrow(fra), 35, replace = T)], trim = 0.05)
```

```
## [1] 309
```

```
mean(Ordered$HSClass[sample(1:nrow(fra), 35, replace = T)], trim = 0.05)
```

```
## [1] 370.9091
```