RWorksheet_Puany#3b

#1

```
#a.
Respondents <- c(1:20)
Sex <- c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2,1,2)
Fathers_Occupation <- c(1,3,3,3,1,2,3,1,1,1,3,2,1,3,3,1,3,1,2,1)
Persons_at_Home <- c(5,7,3,8,5,9,6,7,8,4,7,5,4,7,8,8,3,11,7,6)
Siblings_at_School <- c(6,4,4,1,2,1,5,3,1,2,3,2,5,5,2,1,2,5,3,2)
Types_of_Houses <- c(1,2,3,1,1,3,3,1,2,3,2,3,2,2,3,3,3,3,3,2)

dats <- data.frame(Respondents, Sex, Fathers_Occupation, Persons_at_Home, Siblings_at_School, Types_of_

dats
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1            1   2                  1               5                  6
## 2            2   2                  3               7                  4
## 3            3   1                  3               3                  4
## 4            4   2                  3               8                  1
## 5            5   2                  1               5                  2
## 6            6   2                  2               9                  1
## 7            7   2                  3               6                  5
## 8            8   2                  1               7                  3
## 9            9   2                  1               8                  1
## 10          10   2                  1               4                  2
## 11          11   1                  3               7                  3
## 12          12   2                  2               5                  2
## 13          13   2                  1               4                  5
## 14          14   2                  3               7                  5
## 15          15   2                  3               8                  2
## 16          16   2                  1               8                  1
## 17          17   2                  3               3                  2
## 18          18   2                  1              11                  5
## 19          19   1                  2               7                  3
## 20          20   2                  1               6                  2
##    Types_of_Houses
## 1                1
## 2                2
## 3                3
## 4                1
## 5                1
## 6                3
## 7                3
## 8                1
## 9                2
## 10               3
## 11               2
```

1

```
## 12                3
## 13                2
## 14                2
## 15                3
## 16                3
## 17                3
## 18                3
## 19                3
## 20                2
```

```r
#b
str(dats)
```

```
## 'data.frame':    20 obs. of  6 variables:
##  $ Respondents       : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Sex               : num  2 2 1 2 2 2 2 2 2 2 ...
##  $ Fathers_Occupation: num  1 3 3 3 1 2 3 1 1 1 ...
##  $ Persons_at_Home   : num  5 7 3 8 5 9 6 7 8 4 ...
##  $ Siblings_at_School: num  6 4 4 1 2 1 5 3 1 2 ...
##  $ Types_of_Houses   : num  1 2 3 1 1 3 3 1 2 3 ...
```

```r
summary(dats)
```

```
##   Respondents          Sex        Fathers_Occupation Persons_at_Home
##  Min.   : 1.00   Min.   :1.00   Min.   :1.00       Min.   : 3.0
##  1st Qu.: 5.75   1st Qu.:2.00   1st Qu.:1.00       1st Qu.: 5.0
##  Median :10.50   Median :2.00   Median :2.00       Median : 7.0
##  Mean   :10.50   Mean   :1.85   Mean   :1.95       Mean   : 6.4
##  3rd Qu.:15.25   3rd Qu.:2.00   3rd Qu.:3.00       3rd Qu.: 8.0
##  Max.   :20.00   Max.   :2.00   Max.   :3.00       Max.   :11.0
##  Siblings_at_School Types_of_Houses
##  Min.   :1.00       Min.   :1.0
##  1st Qu.:2.00       1st Qu.:2.0
##  Median :2.50       Median :2.5
##  Mean   :2.95       Mean   :2.3
##  3rd Qu.:4.25       3rd Qu.:3.0
##  Max.   :6.00       Max.   :3.0
```

```r
#c
mean(dats$Siblings_at_School)
```

```
## [1] 2.95
```

No it's not

```r
#d
dats[1:2, ]
```

```
##   Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1           1   2                  1               5                  6
## 2           2   2                  3               7                  4
##   Types_of_Houses
## 1               1
## 2               2
```

```r
#e
dats[c(3,5), c(2,4)]
```

```
##   Sex Persons_at_Home
```

```
## 3     1                 3
## 5     2                 5
```

```r
#f
types_houses <- dats$Types_of_Houses
types_houses
```

```
##  [1] 1 2 3 1 1 3 3 1 2 3 2 3 2 2 3 3 3 3 3 2
```

```r
#g
subset(dats, Sex == 1 & Fathers_Occupation == 1)
```

```
## [1] Respondents        Sex                 Fathers_Occupation Persons_at_Home
## [5] Siblings_at_School Types_of_Houses
## <0 rows> (or 0-length row.names)
```

There are no male respondents whose father is a farmer.

```r
#h
subset(dats, Sex == 2 & Siblings_at_School >= 5)
```

```
##     Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1             1   2                  1               5                  6
## 7             7   2                  3               6                  5
## 13           13   2                  1               4                  5
## 14           14   2                  3               7                  5
## 18           18   2                  1              11                  5
##     Types_of_Houses
## 1                 1
## 7                 3
## 13                2
## 14                2
## 18                3
```

#2

```r
df = data.frame(Ints=integer(),

Doubles=double(), Characters=character(),
Logicals=logical(),
Factors=factor(),
stringsAsFactors=FALSE)

print("Structure of the empty dataframe:")
```

```
## [1] "Structure of the empty dataframe:"
```

```r
print(str(df))
```

```
## 'data.frame':    0 obs. of  5 variables:
##  $ Ints      : int
##  $ Doubles   : num
##  $ Characters: chr
##  $ Logicals  : logi
##  $ Factors   : Factor w/ 0 levels:
## NULL
```

The structure of the data frame is empty, 0 observation/rows and 5 variable/columns with different data
types but has 0 data.

#3

```r
Respondents <- c(1:10)
Sex <- c('Male', 'Female', 'Female', 'Male', 'Male', 'Female', 'Female', 'Male', 'Female', 'Male')
Fathers_Occupation <- c(1,2,3,3,1,2,2,3,1,3)
Persons_at_Home <- c(5,7,3,8,6,4,4,2,11,6)
Siblings_at_School <- c(2,3,0,5,2,3,1,2,6,2)
Types_of_houses <- c('Wood','Concrete','Concrete','Wood', 'Semi-concrete', 'Semi-concrete', 'Wood', 'Se

HouseholdData <- data.frame(Respondents, Sex, Fathers_Occupation, Persons_at_Home, Siblings_at_School,

write.csv(HouseholdData, file = "HouseholdData.csv", row.names = FALSE)
```

#a.
```r
datf <- read.csv("HouseholdData.csv")
datf
```

```
##    Respondents    Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1            1   Male                  1               5                  2
## 2            2 Female                  2               7                  3
## 3            3 Female                  3               3                  0
## 4            4   Male                  3               8                  5
## 5            5   Male                  1               6                  2
## 6            6 Female                  2               4                  3
## 7            7 Female                  2               4                  1
## 8            8   Male                  3               2                  2
## 9            9 Female                  1              11                  6
## 10          10   Male                  3               6                  2
##    Types_of_houses
## 1             Wood
## 2         Concrete
## 3         Concrete
## 4             Wood
## 5    Semi-concrete
## 6    Semi-concrete
## 7             Wood
## 8    Semi-concrete
## 9    Semi-concrete
## 10        Concrete
```

#b.
```r
datf$Sex <- factor(datf$Sex, levels = c("Male", "Female"), labels = c(1, 2))

datf$Sex <- as.integer(as.character(datf$Sex))
datf$Sex
```

```
##  [1] 1 2 2 1 1 2 2 1 2 1
```

#c.
```r
datf$Types_of_houses <- factor(datf$Types_of_houses, levels = c("Wood", "Concrete", "Semi-concrete"), la

datf$Types_of_houses <- as.integer(as.character(datf$Types_of_houses))

datf$Types_of_houses
```

```
##  [1] 1 2 2 1 3 3 1 3 3 2
```

```
#d.
datf$Fathers_Occupation <- factor(datf$Fathers_Occupation, levels = c(1, 2, 3), labels = c("Farmer", "D:

print(datf$Fathers_Occupation)
```

```
##  [1] Farmer Driver Others Others Farmer Driver Driver Others Farmer Others
## Levels: Farmer Driver Others
```

```
#e.
subset(datf, Sex == 2 & Fathers_Occupation == "Driver")
```

```
##   Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 2           2   2             Driver               7                  3
## 6           6   2             Driver               4                  3
## 7           7   2             Driver               4                  1
##   Types_of_houses
## 2               2
## 6               3
## 7               1
```

```
#e.
subset(datf, Siblings_at_School >= 5)
```

```
##   Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 4           4   1             Others               8                  5
## 9           9   2             Farmer              11                  6
##   Types_of_houses
## 4               1
## 9               3
```

#14 Over the span of July 14–21, 2020, Twitter users were mostly sharing negative sentiments. Also, tweet activity dipped on July 20 across all sentiment types. This pattern may reflect public reaction to key events, with emotional intensity peaking on July 15 and July 21, and diping on July 20.