

RWorksheet3b

Abdul Azim Bansara

2023-10-11

#1a

```
household <- data.frame(  
  Respondents = c(1:20),  
  Sex = c("Female","Female","Male","Female","Female","Female","Female","Female","Female","Female","Female","Male",  
  FatherOccupation = c("Farmer","Others","Others","Others","Farmer","Driver","Others","Farmer","Farmer",  
  PersonatHome = c(5,7,3,8,5,9,6,7,8,4,7,5,4,7,8,8,3,11,7,6),  
  Siblingsatschool = c(6,4,4,1,2,1,5,3,1,2,3,2,5,5,2,1,2,5,3,2),  
  Typeshouse = c("Wood","Semi-Concrete","Concrete","Wood","Wood","Concrete","Concrete","Wood","Semi-Concrete",  
)  
household
```

##	Respondents	Sex	FatherOccupation	PersonatHome	Siblingsatschool
## 1	1	Female	Farmer	5	6
## 2	2	Female	Others	7	4
## 3	3	Male	Others	3	4
## 4	4	Female	Others	8	1
## 5	5	Female	Farmer	5	2
## 6	6	Female	Driver	9	1
## 7	7	Female	Others	6	5
## 8	8	Female	Farmer	7	3
## 9	9	Female	Farmer	8	1
## 10	10	Female	Farmer	4	2
## 11	11	Male	Others	7	3
## 12	12	Female	Driver	5	2
## 13	13	Female	Farmer	4	5
## 14	14	Female	Others	7	5
## 15	15	Female	Others	8	2
## 16	16	Female	Farmer	8	1
## 17	17	Female	Others	3	2
## 18	18	Female	Farmer	11	5
## 19	19	Male	Driver	7	3
## 20	20	Female	Farmer	6	2
##	Typeshouse				
## 1	Wood				
## 2	Semi-Concrete				
## 3	Concrete				
## 4	Wood				
## 5	Wood				
## 6	Concrete				
## 7	Concrete				
## 8	Wood				

```
## 9 Semi-Concrete
## 10 Concrete
## 11 Semi-Concrete
## 12 Concrete
## 13 Semi-Concrete
## 14 Semi-Concrete
## 15 Concrete
## 16 Concrete
## 17 Concrete
## 18 Concrete
## 19 Concrete
## 20 Semi-Concrete
```

#1b Describe the data. Get the structure or the summary of the data

```
str(household)
```

```
## 'data.frame': 20 obs. of 6 variables:
## $ Respondents : int 1 2 3 4 5 6 7 8 9 10 ...
## $ Sex : chr "Female" "Female" "Male" "Female" ...
## $ FatherOccupation: chr "Farmer" "Others" "Others" "Others" ...
## $ PersonatHome : num 5 7 3 8 5 9 6 7 8 4 ...
## $ Siblingsatschool: num 6 4 4 1 2 1 5 3 1 2 ...
## $ Typeshouse : chr "Wood" "Semi-Concrete" "Concrete" "Wood" ...
```

#1c

```
mean_siblings <- mean(household$Siblingsatschool)
mean_siblings
```

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

```
mean_siblings == 5
```

```
## [1] FALSE
```

#No because the mean is 2.95

#1d

```
first2rows <- household[1:2, ]
first2rows
```

```
## Respondents Sex FatherOccupation PersonatHome Siblingsatschool
## 1 1 Female Farmer 5 6
## 2 2 Female Others 7 4
## Typeshouse
## 1 Wood
## 2 Semi-Concrete
```

#1e

```
extract35 <- household[c(3,5),c(2,4)]
extract35
```

```
## Sex PersonatHome
## 3 Male 3
## 5 Female 5
```

#1f

```
types_houses <- household$Typeshouse
types_houses
```

```
## [1] "Wood"          "Semi-Concrete" "Concrete"      "Wood"
## [5] "Wood"          "Concrete"      "Concrete"      "Wood"
## [9] "Semi-Concrete" "Concrete"      "Semi-Concrete" "Concrete"
## [13] "Semi-Concrete" "Semi-Concrete" "Concrete"      "Concrete"
## [17] "Concrete"      "Concrete"      "Concrete"      "Semi-Concrete"

#1g
malefarmers <- household[household$Sex == "Male" & household$FatherOccupation == "Farmer", ]
malefarmers

## [1] Respondents      Sex                FatherOccupation PersonatHome
## [5] Siblingsatschool Typeshouse
## <0 rows> (or 0-length row.names)

#there are no male respondents that their father's occupation is a farmer

#1h
femalegor5s <- household[household$Sex == "Female" & household$Siblingsatschool >= 5, ]
femalegor5s

##      Respondents      Sex FatherOccupation PersonatHome Siblingsatschool
## 1              1 Female          Farmer             5             6
## 7              7 Female          Others             6             5
## 13             13 Female          Farmer             4             5
## 14             14 Female          Others             7             5
## 18             18 Female          Farmer            11             5
##      Typeshouse
## 1              Wood
## 7              Concrete
## 13 Semi-Concrete
## 14 Semi-Concrete
## 18              Concrete

#2
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:"

print(str(df))

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

#2a

#the data frame is empty because there is 0 object in the output

#3

```
HouseholdData <- data.frame(
  Respondents = c(1:10),
  Sex = c("Male", "Female", "Female", "Male", "Male", "Female", "Female", "Male", "Female", "Male"),
  FatherOccupation = c("Farmer", "Driver", "Others", "Others", "Farmer", "Driver", "Driver", "Others", "Farmer", "Driver"),
  PersonatHome = c(5, 7, 3, 8, 6, 4, 4, 2, 11, 6),
  Siblingsatschool = c(2, 3, 0, 5, 2, 3, 1, 2, 6, 2),
  Typeshouse = c("Wood", "Congrete", "Congrete", "Wood", "Semi-Congrete", "Semi-Congrete", "Wood", "Semi-Congrete", "Semi-Congrete", "Wood")
)
```

```
HouseholdData

##      Respondents      Sex FatherOccupation PersonatHome Siblingsatschool
## 1             1    Male          Farmer             5           2
## 2             2  Female          Driver             7           3
## 3             3  Female          Others             3           0
## 4             4    Male          Others             8           5
## 5             5    Male          Farmer             6           2
## 6             6  Female          Driver             4           3
## 7             7  Female          Driver             4           1
## 8             8    Male          Others             2           2
## 9             9  Female          Farmer            11           6
## 10            10    Male          Others             6           2
##      Typeshouse
## 1            Wood
## 2          Congrete
## 3          Congrete
## 4            Wood
## 5    Semi-Congrete
## 6    Semi-Congrete
## 7            Wood
## 8    Semi-Congrete
## 9    Semi-Congrete
## 10          Congrete
```

#create a.csv file

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

#3a

```
imported_household <- read.csv("HouseholdData.csv")
imported_household
```

```
##      Respondents      Sex FatherOccupation PersonatHome Siblingsatschool
## 1             1    Male          Farmer             5           2
## 2             2  Female          Driver             7           3
## 3             3  Female          Others             3           0
## 4             4    Male          Others             8           5
## 5             5    Male          Farmer             6           2
## 6             6  Female          Driver             4           3
## 7             7  Female          Driver             4           1
## 8             8    Male          Others             2           2
## 9             9  Female          Farmer            11           6
## 10            10    Male          Others             6           2
```

```
##      Typeshouse
## 1      Wood
## 2      Congrete
## 3      Congrete
## 4      Wood
## 5 Semi-Congrete
## 6 Semi-Congrete
## 7      Wood
## 8 Semi-Congrete
## 9 Semi-Congrete
## 10     Congrete
```

#3b

```
imported_household$Sex <- factor(imported_household$Sex, levels = c("Male", "Female"))
imported_household$Sex <- as.integer(imported_household$Sex)
```

#3c

```
imported_household$Typeshouse <- factor(imported_household$Typeshouse)
imported_household$Typeshouse <- as.integer(factor(imported_household$Typeshouse, levels = c("Wood", "C
```

#3d

```
imported_household$FatherOccupation <- as.integer(factor(imported_household$FatherOccupation, levels = c
```

#3e

```
femalerespo <- imported_household[imported_household$Sex == 2 & imported_household$FatherOccupation == 1, ]
femalerespo
```

```
##  Respondents Sex FatherOccupation PersonatHome Siblingsatschool Typeshouse
## 2           2   2             2           7           3           2
## 6           6   2             2           4           3           3
## 7           7   2             2           4           1           1
```

#3f

```
greaterthan5 <- imported_household[imported_household$Siblingsatschool>=5, ]
greaterthan5
```

```
##  Respondents Sex FatherOccupation PersonatHome Siblingsatschool Typeshouse
## 4           4   1             3           8           5           1
## 9           9   2             1          11           6           3
```

#4 Interpret the graph.

*# on July 14, the negative sentiments were more than the positive and neutral tweets.
on July 15, the negative sentiments were 4000 count and the other sentiment increases too but not high
on July 17, the negative sentiments decreases as well as the other sentiments.
on July 18, the negative sentiments are the same as yesterday and the other sentiments is slightly in
on July 20, all the sentiments decreases
on July 21, the negative sentiments highly increases again but not the same in the July 15th and the*