

Worksheet 3b

Kylene Joy Yanguas

2022-11-13

#1. Create a data frame using the table below.

a. Write the codes.

```
Respondents <- c(1:20)
Sex <-c(2,2,1,2,2,2,2,2,2,1,2,2,2,2,2,2,1,2)
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)

data_frame <- data.frame(Respondents, Sex, Fathers_Occupation,Persons_at_home,
                          Siblings_at_school,Types_of_houses)
data_frame
```

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

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

```
summary(data_frame)
```

```
## 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
```

#c. Is the mean number of siblings attending is 5?

No, the mean is 2.95

*#d. Extract the 1st two rows and then all the columns using the subsetting functions.
#Write the codes and its output.*

```
sub1 <- subset(data_frame[1:2, 1:6, drop = FALSE])
sub1
```

```
## 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
```

*#e. Extract 3rd and 5th row with 2nd and 4th column. Write the codes and its
#result.*

```
sub2 <- subset(data_frame[c(3,5),c(2,4)])
sub2
```

```
## Sex Persons_at_home
## 3 1      3
```

```
## 5      2      5

#f. Select the variable types of houses then store the vector that results as types_houses.
#Write the codes.
sub3 <- subset(data_frame[c(1:20), c(2,6)])

type_houses <- sub3

#g. Select only all Males respondent that their father occupation was farmer. Write
#the codes and its output.

sub4 <- subset(data_frame[c(3,11),c(2,3)])
sub4

##      Sex Fathers_Occupation
## 3      1      3
## 11     1      3

#h. Select only all females respondent that have greater than or equal to 5 number of siblings attendin

sub5 <- subset(data_frame[c(1:20), c(2,5)])
female_resp <- sub5[data_frame$Siblingsatschool >= 5,]
female_resp

## [1] Sex      Siblings_at_school
## <0 rows> (or 0-length row.names)

#2. Write a R program to create an empty data frame. Using the following codes:
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

#a. Describe the results.
#There is zero observations but has a 5 variables. It shows
#that there is no data available in table.

#3. Interpret the graph.
#The data show that the counts of negative sentiments is the highest from July
#14 to 21, 2020.
```