

# Work Sheet 3B

Neil Francis N. Navarro

2022-11-11

## **Worksheet-3b in R**

*Instructions:*

**Use RStudio or the RStudio Cloud accomplish this worksheet.**

*Save the R script as RWorksheet\_lastname 3b.R.*

*On your own GitHub repository, push the R script, the Rmd file, as well as this pdf worksheet to the repo you have created before.*

*Do not forget to comment your Git repo on our VLE*

*Accomplish this worksheet by answering the questions being asked and writing the code manually.*

`install.packages("dplyr") library(dplyr) library(tidyverse)`

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

**a. Write the codes.**

```
Respondents <- c(seq(1,20))
Sex <- c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,1,2)
FathersOccupation <- c(1,3,3,3,1,2,3,1,1,1,3,2,1,3,3,1,3,1,2,1)
Personsathome <- 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)
Typesofhouses <- 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,FathersOccupation,Personsathome,
Siblingsatschool,Typesofhouses)
```

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

```
summary(data_frame)
```

```
## Respondents      Sex      FathersOccupation Personsathome
## 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
## Siblingsatschool Typesofhouses
## 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?

“Answer: No”

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

```
c1 <- subset(data_frame[1:2, 1:6, drop = FALSE])
c1
```

```
## Respondents Sex FathersOccupation Personsathome Siblingsatschool
## 1          1  2              1          5          6
## 2          2  2              3          7          4
## Typesofhouses
## 1          1
## 2          2
```

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

```
c2 <- subset(data_frame[c(3,5),c(2,4)])
c2
```

```
## Sex Personsathome
## 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.

```
c3 <- data_frame[c(6)]
type_houses <- c3
```

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

```
c22 <- subset(data_frame[c(3,11),c(2,3)])
c22
```

```
## Sex FathersOccupation
## 3  1              3
## 11 1              3
```

h. Select only all females respondent that have greater than or equal to 5 number of siblings attending school. Write the codes and its outputs

```
c5 <- subset(data_frame[c(1:20), c(2,5)])
girl <- c5[data_frame$Siblingsatschool >= 5,]
girl
```

```
##      Sex Siblingsatschool
## 1      2                6
## 7      2                5
## 13     2                5
## 14     2                5
## 18     2                5
```

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

```
dtframe = 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(dtframe))
```

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