

RWorksheet_Garrido-3b

Angelo A. Garrido

2025-11-24

1. Create a data frame using the table below

```
#a. Write the codes.
Respondents <- 1:20
Sex = c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,1,2)
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)

respondents_data <- data.frame(
  Respondents,
  Sex,
  Fathers_Occupation,
  Persons_at_Home,
  Siblings_at_School,
  Types_of_Houses
)
```

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

```
respondents_data
```

##	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
```

```
str(respondents_data)
```

```
## '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 ...
```

```
summary(respondents_data)
```

```
##   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?
```

```
mean(Siblings_at_School)
```

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

```
#2.95, not 5
```

```
#d. Extract the 1st two rows and then all the columns using the subsetting functions.
respondents_data[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
```

```
respondents_data
```

```
## 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
```

```
#e. Extract 3rd and 5th row with 2nd and 4th column.
respondents_data[c(3,5), c(2,4)]
```

```
## Sex Persons_at_Home
## 3 1 3
## 5 2 5
```

```
respondents_data
```

```
## 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
```

```
#f. Select the variable types of houses then store the vector that results as types_houses.
types_houses <- respondents_data$Types_of_Houses
types_houses
```

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

#g. Select only all Males respondent that their father occupation was farmer.

```
all_Males <- subset(respondents_data, Sex == 1 & Fathers_Occupation == 1)
all_Males
```

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

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

```
all_Females <- subset(respondents_data, Sex == 2 & Siblings_at_School >= 5)
all_Females
```

```
##      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. 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
```

3. Create a .csv file of this. Save it as HouseholdData.csv

```
Respondents <- 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", "Congrete", "Congrete", "Wood", "Semi-congrete", "Semi-congrete",
```

```

"Wood","Semi-concrete","Semi-concrete","Congrete")

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

HouseholdData

##      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             Congrete
## 3             Congrete
## 4             Wood
## 5      Semi-concrete
## 6      Semi-concrete
## 7             Wood
## 8      Semi-concrete
## 9      Semi-concrete
## 10            Congrete

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

data_set1 <- read.csv("HouseholdData.csv", header = TRUE)
data_set1

```

```

##      X Respondents      Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1    1             1    Male                1             5             2
## 2    2             2 Female                2             7             3
## 3    3             3 Female                3             3             0
## 4    4             4    Male                3             8             5
## 5    5             5    Male                1             6             2
## 6    6             6 Female                2             4             3
## 7    7             7 Female                2             4             1
## 8    8             8    Male                3             2             2
## 9    9             9 Female                1            11             6
## 10  10            10    Male                3             6             2
##      Types_of_Houses
## 1             Wood

```

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

#b. Convert the Sex into factor using factor() function and change it into integer. [Legend: #Male = 1 and Female = 2]. Write the R codes and its output.

```
HouseholdData$Sex <- factor(HouseholdData$Sex, levels = c("Male", "Female"), labels = c(1, 2))
```

#c. Convert the Type of Houses into factor and change it into integer. [Legend: Wood = 1; Congrete = 2; Semi-Congrete = 3]. Write the R codes and its output.

```
HouseholdData$Types_of_Houses <- factor(HouseholdData$Types_of_Houses, levels = c("Wood", "Congrete", "Semi-Congrete"), labels = c(1, 2, 3))
```

```
HouseholdData$Types_of_Houses
```

```
## [1] 1 2 2 1 3 3 1 3 3 2
## Levels: 1 2 3
```

#d. On father's occupation, factor it as Farmer = 1; Driver = 2; and Others = 3. What is the R code and its output?

```
HouseholdData$Fathers_Occupation <- factor(HouseholdData$Fathers_Occupation, levels = c(1, 2, 3), labels = c("Farmer", "Driver", "Others"))
```

```
HouseholdData$Fathers_Occupation
```

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

#e. Select only all females respondent that has a father whose occupation is driver. Write the codes and its output.

```
female_driver <- subset(HouseholdData, Sex == 2 & Fathers_Occupation == "Driver")
female_driver
```

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

#f. Select the respondents that have greater than or equal to 5 number of siblings attending school. Write the codes and its output.

```
many_siblings <- subset(HouseholdData, Siblings_at_School >= 5)
many_siblings
```

```
## Respondents      Sex      Fathers_Occupation Persons_at_Home Siblings_at_School
## 4          4   Male      Others              8              5
## 9          9 Female      Farmer              11              6
## Types_of_Houses
## 4          1
## 9          3
```

```
#e. Extract 3rd and 5th row with 2nd and 4th column.
respondents_data[c(3,5), c(2,4)]
```

```
## Sex Persons_at_Home
## 3 1 3
## 5 2 5
```

```
respondents_data
```

```
## 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
```

```
#f. Select the variable types of houses then store the vector that results as types_houses.
types_houses <- respondents_data$Types_of_Houses
types_houses
```



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

#g. Select only all Males respondent that their father occupation was farmer.

```
all_Males <- subset(respondents_data, Sex == 1 & Fathers_Occupation == 1)
all_Males
```

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

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

```
all_Females <- subset(respondents_data, Sex == 2 & Siblings_at_School >= 5)
all_Females
```

```
##      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. 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
```

3. Create a .csv file of this. Save it as HouseholdData.csv

```
Respondents <- 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", "Congrete", "Congrete", "Wood", "Semi-congrete", "Semi-congrete",
```

```

"Wood", "Semi-concrete", "Semi-concrete", "Congrete")

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

```

HouseholdData

```

##      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             Congrete
## 3             Congrete
## 4             Wood
## 5      Semi-concrete
## 6      Semi-concrete
## 7             Wood
## 8      Semi-concrete
## 9      Semi-concrete
## 10            Congrete

```

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

```

data_set1 <- read.csv("HouseholdData.csv", header = TRUE)
data_set1

```

```

##      X Respondents      Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1    1             1    Male                1             5             2
## 2    2             2 Female                2             7             3
## 3    3             3 Female                3             3             0
## 4    4             4    Male                3             8             5
## 5    5             5    Male                1             6             2
## 6    6             6 Female                2             4             3
## 7    7             7 Female                2             4             1
## 8    8             8    Male                3             2             2
## 9    9             9 Female                1            11             6
## 10  10            10    Male                3             6             2
##      Types_of_Houses
## 1             Wood

```

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

#b. Convert the Sex into factor using factor() function and change it into integer. [Legend: #Male = 1 and Female = 2]. Write the R codes and its output.

```
HouseholdData$Sex <- factor(HouseholdData$Sex, levels = c("Male", "Female"), labels = c(1, 2))
```

#c. Convert the Type of Houses into factor and change it into integer. [Legend: Wood = 1; Congrete = 2; Semi-Congrete = 3]. Write the R codes and its output.

```
HouseholdData$Types_of_Houses <- factor(HouseholdData$Types_of_Houses, levels = c("Wood", "Congrete", "Semi-Congrete"), labels = c(1, 2, 3))
```

```
HouseholdData$Types_of_Houses
```

```
## [1] 1 2 2 1 3 3 1 3 3 2
## Levels: 1 2 3
```

#d. On father's occupation, factor it as Farmer = 1; Driver = 2; and Others = 3. What is the R code and its output?

```
HouseholdData$Fathers_Occupation <- factor(HouseholdData$Fathers_Occupation, levels = c(1, 2, 3), labels = c("Farmer", "Driver", "Others"))
```

```
HouseholdData$Fathers_Occupation
```

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

#e. Select only all females respondent that has a father whose occupation is driver. Write the codes and its output.

```
female_driver <- subset(HouseholdData, Sex == 2 & Fathers_Occupation == "Driver")
female_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
```

#f. Select the respondents that have greater than or equal to 5 number of siblings attending school. Write the codes and its output.

```
many_siblings <- subset(HouseholdData, Siblings_at_School >= 5)
many_siblings
```

```
## 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
```

4. Interpret the graph.

#Based on the results of the graph, negative opinion dominated the use of Twitter between July 14-21, 2020