RWorksheet_Caoyonan#3b.Rmd

Leanie G. Caoyonan

2025-10-13

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

#a. Write the codes.
household <- data.frame(
  Sex = c("Male", "Female", "Male", "Female",    "Male"),

  Types_of_Houses = c("Wood", "Concrete", "Semi-Concrete", "Wood", "Concrete"),

  Fathers_Occupation = c("Farmer", "Driver", "Others", "Farmer", "Driver"),

  Num_Siblings_Attending = c(4, 6, 5, 3, 7)
)


#b. Describe the data. Get the structure or the summary of the data
str(household)
```

```
## 'data.frame':    5 obs. of  4 variables:
##  $ Sex                   : chr  "Male" "Female" "Male" "Female" ...
##  $ Types_of_Houses       : chr  "Wood" "Concrete" "Semi-Concrete" "Wood" ...
##  $ Fathers_Occupation    : chr  "Farmer" "Driver" "Others" "Farmer" ...
##  $ Num_Siblings_Attending: num  4 6 5 3 7
```

```
summary(household)
```

```
##      Sex            Types_of_Houses    Fathers_Occupation
##  Length:5           Length:5           Length:5
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##  Num_Siblings_Attending
##  Min.   :3
##  1st Qu.:4
##  Median :5
##  Mean   :5
##  3rd Qu.:6
##  Max.   :7
```

```
#c. Is the mean number of siblings attending is 5?
mean(household$Num_Siblings_Attending) == 5
```

```
## [1] TRUE
```

```r
#d. Extract the 1st two rows and then all the columns using the sub-setting functions. Write the codes
household[1:2, ]
```

```
##      Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 1   Male            Wood             Farmer                     4
## 2 Female        Concrete             Driver                     6
```

```r
#e. Extract 3rd and 5th row with 2nd and 4th column. Write the codes and its result.
household[c(3,5), c(2,4)]
```

```
##   Types_of_Houses Num_Siblings_Attending
## 3   Semi-Concrete                     5
## 5        Concrete                     7
```

```r
#Output:
#       Types_of_Houses
#     3  Semi_Concrete
#     5  Concrete


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

```
## [1] "Wood"          "Concrete"      "Semi-Concrete" "Wood"
## [5] "Concrete"
```

```r
#Output:
#[1] "Wood"          "Concrete"
#[3] "Semi-Concrete" "Wood"
#[5] "Concrete"


#g. Select only all Males respondent that their father occupation was farmer. Write the codes and its o
subset(household, Sex == "Male" & Fathers_Occupation == "Farmer")
```

```
##    Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 1 Male            Wood             Farmer                     4
```

```r
#Output:

#     Sex        Types_of_Houses
#  1  Male      Wood


#h. Select only all females respondent that have greater than or equal to 5 number of siblings attendin
subset(household, Sex == "Female" & Num_Siblings_Attending >= 5)
```

```
##      Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 2 Female        Concrete             Driver                     6
```

```r
#Output:
#       Sex       Types_of_Houses
#    2  Female    Concrete


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

```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
str(df)
```

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

```r
#a. Describe the results.
#Describe: The result shows that the empty data frame has 0 rows and 5 columns with different data type


#3. Create a .csv file of this. Save it as HouseholdData.csv
write.csv(household, "HouseholdData.csv", row.names = FALSE)


#a. Import the csv file into the R environment. Write the codes.
HouseholdData <- read.csv("HouseholdData.csv")
HouseholdData
```

```
##      Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 1   Male            Wood             Farmer                     4
## 2 Female        Concrete             Driver                     6
## 3   Male   Semi-Concrete             Others                     5
## 4 Female            Wood             Farmer                     3
## 5   Male        Concrete             Driver                     7
```

```r
#b. Convert the Sex into factor using factor() function and change it into integer.[Legend: Male = 1 an
HouseholdData$Sex <- factor(HouseholdData$Sex, levels = c("Male", "Female"), labels = c(1, 2))
HouseholdData
```

```
##   Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 1   1            Wood             Farmer                     4
## 2   2         Concrete             Driver                     6
## 3   1    Semi-Concrete             Others                     5
## 4   2            Wood             Farmer                     3
## 5   1         Concrete             Driver                     7
```

```r
#Output:


#c. Convert the Type of Houses into factor and change it into integer. [Legend: Wood = 1; Congrete = 2;
HouseholdData$Types_of_Houses <- factor(
```

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

```
##   Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 1   1               1             Farmer                      4
## 2   2               2             Driver                      6
## 3   1               3             Others                      5
## 4   2               1             Farmer                      3
## 5   1               2             Driver                      7
```

```
#Output:


#d. On father's occupation, factor it as Farmer = 1; Driver = 2; and Others = 3. What is the R code and
HouseholdData$Fathers_Occupation <- factor(
  HouseholdData$Fathers_Occupation,
  levels = c("Farmer", "Driver", "Others"),
  labels = c(1, 2, 3)
)
HouseholdData
```

```
##   Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 1   1               1                 1                      4
## 2   2               2                 2                      6
## 3   1               3                 3                      5
## 4   2               1                 1                      3
## 5   1               2                 2                      7
```

```
#Output:


#e. Select only all females respondent that has a father whose occupation is driver. Write the codes an
subset(HouseholdData, Sex == 2 & Fathers_Occupation == 2)
```

```
##   Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 2   2               2                 2                      6
```

```
#Output:


#f. Select the respondents that have greater than or equal to 5 number of siblings attending school. Wr
subset(HouseholdData, Num_Siblings_Attending >= 5)
```

```
##   Sex Types_of_Houses Fathers_Occupation Num_Siblings_Attending
## 2   2               2                 2                      6
## 3   1               3                 3                      5
## 5   1               2                 2                      7
```

```
#Output:
```