RWorksheet_Anastacio#3b

Kyle Johnsen Anastacio

2025-10-26

1. Create a data frame using the table below.

a. Write the codes.

```r
Respondents <- 1:20

Sex <- c(2,2,1,2,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2,1,2)
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)

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

df
```

```
##    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

```r
str(df)
```

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

```r
summary(df)
```

```
##   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?

```r
mean_siblings <- mean(df$Siblings_at_School)
mean_siblings
```

```
## [1] 2.95
mean_siblings == 5
```

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

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

```
first_two <- df[1:2, ]
first_two
```

```
##   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.

```
df[c(3,5), c(2,4)]
```

```
##   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.

```
types_houses <- df$Types_of_Houses
types_houses
```

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

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

```
male_farmer <- subset(df, Sex == "Male" & Fathers_Occupation == "Farmer")
male_farmer
```

```
## [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. Write the codes and its outputs.

```
female_sib_ge5 <- subset(df, Sex == "Female" & Siblings_at_School >= 5)
female_sib_ge5
```

```
## [1] Respondents        Sex                Fathers_Occupation Persons_at_Home
## [5] Siblings_at_School Types_of_Houses
## <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:") print(str(df))

a. Describe the results.