RWorksheet_Layson#3b

Gilmar Layson

2024-10-07

#1. a.

```
datafr <- data.frame(
  Respondents = c(1:20),
  Sex = c(2, 2, 1, 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, 8, 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)
)
datafr
```

```
##    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              8                  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.

```
str(datafr)
```

```
## '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(datafr)
```

```
##   Respondents          Sex       Fathers_Occupation Persons_at_home
##  Min.   : 1.00   Min.   :1.00   Min.   :1.00       Min.   : 3.00
##  1st Qu.: 5.75   1st Qu.:2.00   1st Qu.:1.00       1st Qu.: 5.00
##  Median :10.50   Median :2.00   Median :2.00       Median : 7.00
##  Mean   :10.50   Mean   :1.85   Mean   :1.95       Mean   : 6.55
##  3rd Qu.:15.25   3rd Qu.:2.00   3rd Qu.:3.00       3rd Qu.: 8.00
##  Max.   :20.00   Max.   :2.00   Max.   :3.00       Max.   :11.00
##  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.

```
mean(datafr[,5])
```

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

#d.

```
first2 <- datafr[1:2, ]
first2
```

```
##   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.
```

row_and_col <- datafr[c(3,5), c(2,4)] row_and_col

#f.

```
types_houses <- datafr[ ,6]
types_houses
```

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

#g.

```
Farm_Male <- subset(datafr, datafr[,2] == 1 & datafr[,3] == 1 )
Farm_Male
```

```
## [1] Respondents        Sex                Fathers_Occupation Persons_at_home
## [5] Siblings_at_school Types_of_houses
## <0 rows> (or 0-length row.names)
```

#h.

```
Female_Res <- subset(datafr, datafr[,2] == 2 & datafr[,5] >= 5)
Female_Res
```

```
##    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.

```
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. #Five columns (Int, Doubles, Characters, Logicals, and Factors) with different data types (integer, numeric, character, logical, and factor) but no observations and make up the empty data frame.

#3 a.

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

```
##    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-congrete
## 6    Semi-congrete
## 7             wood
## 8    Semi-congrete
## 9    Semi-congrete
## 10        Congrete
```

b.

```r
factor(household[,2], levels = c("Male", "Female"), labels = c(1,2))
```

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

c.

```r
factor(household[,6], levels = c("Wood", "Congrete", "Semi-congrete"), labels = c(1,2,3))
```

```
##  [1] 1    2    2    1    3    3    <NA> 3    3    2
## Levels: 1 2 3
```

d.

```r
factor(household[,3], levels = c(1,2,3), labels = c("Farmer", "Driver", "Others"))
```

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

e.

```r
subset(household[,c(2,3)], household[,2] == 2 & household[,3] == "Driver")
```

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

f.

```r
subset(household, household[,5] >= 5)
```

```
##   Respondents    Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 4           4   Male                  3              8                  5
## 9           9 Female                  1             11                  6
##   Types.of.Houses
## 4            Wood
## 9   Semi-congrete
```

#4. #The data table describes the demographics of 10 respondents, including their sex, the jobs held by their fathers, the number of people living at home, the number of siblings enrolled in school, and the kinds of homes they reside in.