RWorksheet_guion#3b

Mikyla Grace Guion

2024-10-02

1.

a.

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

dfOne
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 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
##    TypesOfHouses
## 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(dfOne)
```

```
## '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 ...
##  $ FathersOccupation: num  1 3 3 3 1 2 3 1 1 1 ...
##  $ PersonsAtHome    : num  5 7 3 8 5 9 6 7 8 4 ...
##  $ SiblingsAtSchool : num  6 4 4 1 2 1 5 3 1 2 ...
##  $ TypesOfHouses    : num  1 2 3 1 1 3 3 1 2 3 ...
```

#The structure shows the number of objects and variables in the data frame. It shows the first few contents of the data frame and also the data type of each column.

c.

```
mean(dfOne$SiblingsAtSchool)
```

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

d.

```
subset<- dfOne[1:2, ]
subset
```

```
##   Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 1           1   2                 1             5                6
## 2           2   2                 3             7                4
##   TypesOfHouses
## 1             1
## 2             2
```

e.

```
subSetOne <- dfOne[c(3, 5), c(2, 4)]
subSetOne
```

```
##   Sex PersonsAtHome
## 3   1             3
## 5   2             5
```

f.

```
types.houses <- dfOne$TypesOfHouses
```

g.

```
maleFarmers <- subset(dfOne, Sex == 1 & FathersOccupation == 1)
maleFarmers
```

```
## [1] Respondents       Sex               FathersOccupation PersonsAtHome
## [5] SiblingsAtSchool  TypesOfHouses
## <0 rows> (or 0-length row.names)
```

h.

```
femaleSiblings <- subset(dfOne, Sex == 2 & SiblingsAtSchool >= 5)
femaleSiblings
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 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
##    TypesOfHouses
## 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:"
```

```r
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.

The result shows the structure of the empty data frame. It shows 0 observations and 5 variables. I also shows different data types with no data yet.

3.

a.

```r
HouseHold <- read.csv("HouseholdData.csv")
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
HouseHold$Sex <- factor(HouseHold$Sex, levels = c("Male", "Female"), labels = c(1, 2))
HouseHold$Sex <- as.integer(as.character(HouseHold$Sex))
```

```
HouseHold
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1            1   1                  1               5                  2
## 2            2   2                  2               7                  3
## 3            3   2                  3               3                  0
## 4            4   1                  3               8                  5
## 5            5   1                  1               6                  2
## 6            6   2                  2               4                  3
## 7            7   2                  2               4                  1
## 8            8   1                  3               2                  2
## 9            9   2                  1              11                  6
## 10          10   1                  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
```

c.

```
HouseHold$Types_of_Houses <- factor(HouseHold$Types_of_Houses, levels = c("Wood", "Congrete", "Semi-cong
HouseHold$Types_of_Houses <- as.integer(as.character(HouseHold$Types_of_Houses))
HouseHold
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1            1   1                  1               5                  2
## 2            2   2                  2               7                  3
## 3            3   2                  3               3                  0
## 4            4   1                  3               8                  5
## 5            5   1                  1               6                  2
## 6            6   2                  2               4                  3
## 7            7   2                  2               4                  1
## 8            8   1                  3               2                  2
## 9            9   2                  1              11                  6
## 10          10   1                  3               6                  2
##    Types_of_Houses
## 1                1
## 2                2
## 3                2
## 4                1
## 5                3
## 6                3
## 7                1
## 8                3
## 9                3
## 10               2
```

d.

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

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1            1   1             Farmer               5                  2
## 2            2   2             Driver               7                  3
## 3            3   2             Others               3                  0
## 4            4   1             Others               8                  5
## 5            5   1             Farmer               6                  2
## 6            6   2             Driver               4                  3
## 7            7   2             Driver               4                  1
## 8            8   1             Others               2                  2
## 9            9   2             Farmer              11                  6
## 10          10   1             Others               6                  2
##    Types_of_Houses
## 1                1
## 2                2
## 3                2
## 4                1
## 5                3
## 6                3
## 7                1
## 8                3
## 9                3
## 10               2
```

e.

```
femaleDriver <- subset(HouseHold, Sex == 2 & Fathers_Occupation == "Driver")
femaleDriver
```

```
##   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.

```
SiblingSchool <- subset(HouseHold, Siblings_at_School >= 5)
SiblingSchool
```

```
##   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