RWorksheet_Aposaga#3b

John Philipp Aposaga

2024-10-01

1.

a.

```
dataf <- 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)
)
dataf
```

```
##    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(dataf)
```

```
## '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(dataf)
```

```
##   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(dataf[,5])
```

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

d.

```r
first2 <- dataf[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.

```r
rowncol <- dataf[c(3,5), c(2,4)]
rowncol
```

```
##   Sex Persons_at_home
## 3   1               3
## 5   2               5
```

f.

```r
types_houses <- dataf[,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.

```r
maleFarm <- subset(dataf, dataf[,2] == 1 & dataf[,3] == 1 )
maleFarm
```

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

h.

```r
femaleSchl <- subset(dataf, dataf[,2] == 2 & dataf[,5] >= 5)
femaleSchl
```

```
##    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.

a.

```
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.

a.

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

```
##    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.

```
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.

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

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

d.

```
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.

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

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

f.

```
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
```

g.

```
#The graph highlights how people's feelings in tweets changed over several days,
#with shifts in positive, neutral, and negative sentiments. Negative tweets
#peaked on July 15, likely in response to bad news, while there was an increase
#in positive tweets on July 17, indicating a better mood. By July 20, neutral
#tweets became the most common, suggesting more balanced or informational
#posts during that time.
```