RWorksheet_Lomibao#3b

lomibao

2024-10-03

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, 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)
)
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               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.It shows the summary of the data such as mean, median, max, min and etc.

```
summary(dataf)
```

```
##   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. No its 2.95. You could see it at the data itself. D.

```
subf <- dataf[1:2 , ]
subf
```

```
##   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.

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

```
##   Sex Persons_at_Home
## 3   1               3
## 5   2               5
```

F.

```
types_houses <- dataf[ , "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.

```
mal_res <- subset(dataf, Sex == 1 & Fathers_Occupation == 1)
mal_res
```

```
## [1] Respondents        Sex                Fathers_Occupation Persons_at_Home
## [5] Siblings_at_School Types_of_Houses
## <0 rows> (or 0-length row.names)
```

H.

```
fem_res <- subset(dataf, Sex ==2 & Siblings_at_School >=5)
fem_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.It show what is within the dataframe such as objects, variables and data types.

```
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.

```
HouseholdData <- read.csv("/cloud/project/HouseholdData - Sheet1.csv")
HouseholdData
```

```
##    Respondents    Sex Fathers_Occupation Person_at_home Siblings_at_schools
## 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.

```
HouseholdData$Sex <- factor(HouseholdData$Sex, levels = c("Male","Female"), labels = c(1,2))
HouseholdData
```

```
##    Respondents Sex Fathers_Occupation Person_at_home Siblings_at_schools
## 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.

```
HouseholdData$Types_of_houses <- factor(HouseholdData$Types_of_houses, levels = c("Wood", "Congrete", "
HouseholdData
```

```
##    Respondents Sex Fathers_Occupation Person_at_home Siblings_at_schools
## 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.

```
HouseholdData$Fathers_Occupation <- factor(HouseholdData$Fathers_Occupation, levels = c(1, 2, 3), labels
HouseholdData
```

```
##    Respondents Sex Fathers_Occupation Person_at_home Siblings_at_schools
## 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.

```
female_res <- subset(HouseholdData, Sex==2 & Fathers_Occupation =="Driver")
female_res
```

```
##   Respondents Sex Fathers_Occupation Person_at_home Siblings_at_schools
## 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.

```
ms <- subset(HouseholdData, Respondents & Siblings_at_schools >=  5)
ms
```

```
##   Respondents Sex Fathers_Occupation Person_at_home Siblings_at_schools
## 4           4   1             Others              8                   5
## 9           9   2             Farmer             11                   6
##   Types_of_houses
## 4               1
## 9               3
```

4. The graph show the sentiment of tweets from July 14, 2020 to July 21 ,2020. It shows that negative and neutral Tweets reach their peaks in July 15,2020 While the positive tweets reach their peaks in July 21, 2020.