Rworksheet_Mabalina#3b

Kristian B. Mabalina

2024-10-02

##1 ##A

```
SurveyData <- data.frame(
  RespondentID = c(1:20),
  Gender = c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2, 2, 1,2),
  Fathers_Job = c(1, 3, 3, 3, 1, 2, 3, 1, 1, 1, 3, 2, 1, 3, 3, 1, 3, 1, 2, 1),
  HouseholdSize = c(5, 7, 3, 8, 5, 9, 6, 7, 8, 4, 7, 5, 4, 7, 8, 8, 3, 11, 7, 6),
  SchoolAgeSiblings = c(6, 4, 4, 1, 2, 1, 5, 3, 1, 2, 3, 2, 5, 5, 2, 1, 2, 5, 3, 2),
  HomeType = c(1, 2, 3, 1, 1, 3, 3, 1, 2, 3, 2, 3, 2, 2, 3, 3, 3, 3, 3, 2)
)

SurveyData
```

```
##    RespondentID Gender Fathers_Job HouseholdSize SchoolAgeSiblings HomeType
## 1             1      2           1             5                 6        1
## 2             2      2           3             7                 4        2
## 3             3      1           3             3                 4        3
## 4             4      2           3             8                 1        1
## 5             5      2           1             5                 2        1
## 6             6      2           2             9                 1        3
## 7             7      2           3             6                 5        3
## 8             8      2           1             7                 3        1
## 9             9      2           1             8                 1        2
## 10           10      2           1             4                 2        3
## 11           11      1           3             7                 3        2
## 12           12      2           2             5                 2        3
## 13           13      2           1             4                 5        2
## 14           14      2           3             7                 5        2
## 15           15      2           3             8                 2        3
## 16           16      2           1             8                 1        3
## 17           17      2           3             3                 2        3
## 18           18      2           1            11                 5        3
## 19           19      1           2             7                 3        3
## 20           20      2           1             6                 2        2
```

##B

```
str(SurveyData)
```

```
## 'data.frame':    20 obs. of  6 variables:
##  $ RespondentID     : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Gender           : num  2 2 1 2 2 2 2 2 2 2 ...
##  $ Fathers_Job      : num  1 3 3 3 1 2 3 1 1 1 ...
##  $ HouseholdSize    : num  5 7 3 8 5 9 6 7 8 4 ...
```

```
##  $ SchoolAgeSiblings: num  6 4 4 1 2 1 5 3 1 2 ...
##  $ HomeType         : num  1 2 3 1 1 3 3 1 2 3 ...
```

```
summary(SurveyData)
```

```
##   RespondentID       Gender      Fathers_Job   HouseholdSize  SchoolAgeSiblings
##  Min.   : 1.00   Min.   :1.00   Min.   :1.00   Min.   : 3.0   Min.   :1.00
##  1st Qu.: 5.75   1st Qu.:2.00   1st Qu.:1.00   1st Qu.: 5.0   1st Qu.:2.00
##  Median :10.50   Median :2.00   Median :2.00   Median : 7.0   Median :2.50
##  Mean   :10.50   Mean   :1.85   Mean   :1.95   Mean   : 6.4   Mean   :2.95
##  3rd Qu.:15.25   3rd Qu.:2.00   3rd Qu.:3.00   3rd Qu.: 8.0   3rd Qu.:4.25
##  Max.   :20.00   Max.   :2.00   Max.   :3.00   Max.   :11.0   Max.   :6.00
##     HomeType
##  Min.   :1.0
##  1st Qu.:2.0
##  Median :2.5
##  Mean   :2.3
##  3rd Qu.:3.0
##  Max.   :3.0
```

##C

```
meanSchoolSiblings <- mean(SurveyData$SchoolAgeSiblings)
meanSchoolSiblings
```

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

##D

```
subset_Respondents <- subset(SurveyData, RespondentID <= 2)
subset_Respondents
```

```
##   RespondentID Gender Fathers_Job HouseholdSize SchoolAgeSiblings HomeType
## 1            1      2           1             5                6        1
## 2            2      2           3             7                4        2
```

##E

```
specificData <- SurveyData[c(3,5), c("Gender", "HouseholdSize")]
specificData
```

```
##   Gender HouseholdSize
## 3      1             3
## 5      2             5
```

##F

```
homeTypes <- SurveyData$HomeType
homeTypes
```

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

##G

```
maleFarmers <- SurveyData[SurveyData$Gender == 1 & SurveyData$Fathers_Job == 1, ]
maleFarmers
```

```
## [1] RespondentID      Gender            Fathers_Job       HouseholdSize
## [5] SchoolAgeSiblings HomeType
## <0 rows> (or 0-length row.names)
```

##H

```
femaleManySiblings <- SurveyData[SurveyData$Gender == 2 & SurveyData$SchoolAgeSiblings >= 5, ]
femaleManySiblings
```

```
##    RespondentID Gender Fathers_Job HouseholdSize SchoolAgeSiblings HomeType
## 1             1      2           1             5                 6        1
## 7             7      2           3             6                 5        3
## 13           13      2           1             4                 5        2
## 14           14      2           3             7                 5        2
## 18           18      2           1            11                 5        3
```

##2

```
df_empty <- data.frame(
  Ints = integer(),
  Doubles = double(),
  Strings = character(),
  Booleans = logical(),
  Categories = factor(),
  stringsAsFactors = FALSE
)
print("Structure of the empty data frame:")
```

```
## [1] "Structure of the empty data frame:"
```

```
str(df_empty)
```

```
## 'data.frame':    0 obs. of  5 variables:
##  $ Ints      : int
##  $ Doubles   : num
##  $ Strings   : chr
##  $ Booleans  : logi
##  $ Categories: Factor w/ 0 levels:
```

##3 ##A

```
HouseHold <- read.csv("HouseholdData.csv")
HouseHold
```

```
##    Respondents    Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1            1   Male                 1              5                 5
## 2            2 Female                 2              7                 3
## 3            3 Female                 3              3                 0
## 4            4   Male                 3              8                 8
## 5            5   Male                 1              6                 6
## 6            6 Female                 2              4                 4
## 7            7 Female                 2              4                 4
## 8            8   Male                 3              2                 2
## 9            9 Female                 3             11                 6
## 10          10   Male                 3              6                 6
##    Types.of.Houses
## 1             Wood
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-concrete
## 6    Semi-concrete
```

```
## 7             Wood
## 8    Semi-concrete
## 9    Semi-concrete
## 10        Congrete
```

##B

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

```
##    Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1            1   1                  1               5                  5
## 2            2   2                  2               7                  3
## 3            3   2                  3               3                  0
## 4            4   1                  3               8                  8
## 5            5   1                  1               6                  6
## 6            6   2                  2               4                  4
## 7            7   2                  2               4                  4
## 8            8   1                  3               2                  2
## 9            9   2                  3              11                  6
## 10          10   1                  3               6                  6
##    Types.of.Houses
## 1             Wood
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-concrete
## 6    Semi-concrete
## 7             Wood
## 8    Semi-concrete
## 9    Semi-concrete
## 10        Congrete
```

##C

```
HouseHold$Types.of.Houses <- factor(HouseHold$Types.of.Houses,
                                    levels = c("Wood", "Concrete", "Semi-concrete"),
                                    labels = c(1, 2, 3))
HouseHold$Types.of.Houses <- as.integer(HouseHold$Types.of.Houses)
HouseHold
```

```
##    Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1            1   1                  1               5                  5
## 2            2   2                  2               7                  3
## 3            3   2                  3               3                  0
## 4            4   1                  3               8                  8
## 5            5   1                  1               6                  6
## 6            6   2                  2               4                  4
## 7            7   2                  2               4                  4
## 8            8   1                  3               2                  2
## 9            9   2                  3              11                  6
## 10          10   1                  3               6                  6
##    Types.of.Houses
## 1                1
## 2               NA
## 3               NA
```

```
## 4                1
## 5                3
## 6                3
## 7                1
## 8                3
## 9                3
## 10              NA
```

##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                  5
## 2            2   2             Driver               7                  3
## 3            3   2             Others               3                  0
## 4            4   1             Others               8                  8
## 5            5   1             Farmer               6                  6
## 6            6   2             Driver               4                  4
## 7            7   2             Driver               4                  4
## 8            8   1             Others               2                  2
## 9            9   2             Others              11                  6
## 10          10   1             Others               6                  6
##    Types.of.Houses
## 1                1
## 2               NA
## 3               NA
## 4                1
## 5                3
## 6                3
## 7                1
## 8                3
## 9                3
## 10              NA
```

##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                  4
## 7           7   2             Driver               4                  4
##   Types.of.Houses
## 2              NA
## 6               3
## 7               1
```

##F

```
SiblingSchool <- subset(HouseHold, Siblings.at.School >= 5)
SiblingSchool
```

```
##    Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
```

```
## 1            1   1            Farmer               5                 5
## 4            4   1            Others               8                 8
## 5            5   1            Farmer               6                 6
## 9            9   2            Others              11                 6
## 10          10   1            Others               6                 6
##    Types.of.Houses
## 1                1
## 4                1
## 5                3
## 9                3
## 10              NA
```

##4 The graph shows that between July 14 and July 21, 2020, there were more negative tweets each day than any other type. Positive tweets were the second highest, and there were the least neutral tweets.