

RWorksheet_lastname#3b.Rmd

Jowi Lorenz Nyl Joven BSIT 2-B

2023-11-22

#1. Create a data frame using the table below.

#a.

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
Respondents <- c(seq(1,20))
```

```
Sex <- c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,1,2)
```

```
Father_Occupation <- 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)
```

```
TypeOfHouses <- c(1,2,3,1,1,3,3,1,2,3,2,3,2,2,3,3,3,3,3,2)
```

```
data_display <- data.frame(Respondents, Sex, Father_Occupation, PersonsAtHome, SiblingsAtSchool, TypeOfHouses)
```

```
## Respondents Sex Father_Occupation 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
##   TypeOfHouses
## 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
```

```
#The dataset consists of responses from 20 individuals, providing details about their gender, with 7 ma
summary(data_display)
```

```
##   Respondents      Sex      Father_Occupation PersonsAtHome
##   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
##   SiblingsAtSchool TypeOfHouses
##   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.
```

```
#d.
```

```
data1 <- subset (data_display)[1:2, 2:6, drop=FALSE]
data1
```

```
##   Sex Father_Occupation PersonsAtHome SiblingsAtSchool TypeOfHouses
## 1   2           1           5           6           1
## 2   2           3           7           4           2
```

```
#e.
```

```
data2 <- data_display[c(3,5), c(2,4)]
data2
```

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

```
#f.
types_houses <- data_display[c(6)]
types_houses
```

```
##      TypeOfHouses
## 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
```

```
#g.
selected_data <- data_display %>% select(1:6)
data3 <- selected_data[data_display$Sex == 1,]
data3
```

```
##      Respondents Sex Father_Occupation PersonsAtHome SiblingsAtSchool
## 3              3  1              3              3              4
## 11             11  1              3              7              3
## 19             19  1              2              7              3
##      TypeOfHouses
## 3              3
## 11             2
## 19             3
```

```
#h.
female <- selected_data[data_display$SiblingsAtSchool >= 5,]
female
```

```
##      Respondents Sex Father_Occupation 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
##      TypeOfHouses
## 1              1
```

```
## 7          3
## 13         2
## 14         2
## 18         3
```

#2. Write a R program to create an empty data frame. Using the following codes:

```
df = data.frame(Ids=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:
## $ Ids          : int
## $ Doubles       : num
## $ Characters    : chr
## $ Logicals     : logi
## $ Factors      : Factor w/ 0 levels:
## NULL
```

#3. Create a .csv file of this. Save it as HouseholdData.csv

#a

```
RespondentsNew<-c(1,2,3,4,5,6,7,8,9,10)
SexNew<-c("Male", "Female", "Female", "Male", "Male", "Female", "Female", "Male", "Female", "Male")
FathersOccupationNew<-c(1,2,3,3,1,2,2,3,1,3)
PeAtHomeNew<-c(5,7,3,8,6,4,4,2,11,6)
SibAtSchoolNew<-c(2,3,0,5,2,3,1,2,6,2)
TypesofHousesNew<-c("Wood", "Congrete", "Congrete", "Wood", "Semi-Congrete", "Semi-Congrete", "Wood", "Semi-Congrete", "Semi-Congrete", "Wood")
HouseholdData<-data.frame(
  RespondentsNew,
  SexNew,
  FathersOccupationNew,
  PeAtHomeNew,
  SibAtSchoolNew,
  TypesofHousesNew
)
HouseholdData
```

```
##      RespondentsNew SexNew FathersOccupationNew PeAtHomeNew SibAtSchoolNew
## 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
##      TypesofHousesNew
```

```
## 1      Wood
## 2      Congrete
## 3      Congrete
## 4      Wood
## 5      Semi-Congrete
## 6      Semi-Congrete
## 7      Wood
## 8      Semi-Congrete
## 9      Semi-Congrete
## 10     Congrete
```

```
library(readr)
csv_file <- "HouseholdData.csv"
write.csv(HouseholdData, file = csv_file)
HouseholdData <- read.csv("HouseholdData.csv")
#4
#b
data_display1 <- factor(HouseholdData$SexNew, levels = c("Male" = 1, "Female" = 2))
sex_mapping <- c("Male" = 1, "Female" = 2)
data_display1<-as.integer(sex_mapping[HouseholdData$SexNew])
unique(data_display1)
```

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

```
unique(HouseholdData$SexNew)
```

```
## [1] "Male" "Female"
```

```
#c.
data_display2 <- factor(HouseholdData$TypesofHousesNew, levels = c("Wood" = 1, "Congrete" = 2,"Semi-Congrete" = 3))
sex_mapping2 <- c("Wood" = 1, "Congrete" = 2,"Semi-Congrete" = 3)
data_display2 <- as.integer(sex_mapping2[HouseholdData$TypesofHousesNew])
unique(data_display2)
```

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

```
unique(HouseholdData$TypesofHousesNew)
```

```
## [1] "Wood" "Congrete" "Semi-Congrete"
```

```
#d.
data_display3 <- factor(HouseholdData$FathersOccupationNew, labels=c("Farmer" = 1, "Driver" = 2,"Others" = 3))
sex_mapping3 <- c("Farmer" = 1, "Driver" = 2,"Others" = 3)
data_display3 <- as.integer(sex_mapping3[HouseholdData$FathersOccupationNew])
unique(data_display3)
```

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

```
unique(HouseholdData$FathersOccupationNew)
```

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

```
#e.
selected_data3 <- HouseholdData %>% select(2,3,4)
data4 <- selected_data3[HouseholdData$FathersOccupationNew == 2, ]
data4
```

```
## RespondentsNew SexNew FathersOccupationNew
## 2      2 Female      2
```

```
## 6          6 Female          2
## 7          7 Female          2
```

```
#f.
selected_data3 <- HouseholdData %>% select(2,6)
data4 <- selected_data3[HouseholdData$SibAtSchoolNew >= 5,]
data4
```

```
## RespondentsNew SibAtSchoolNew
## 4          4          5
## 9          9          6
```

```
colnames(HouseholdData) <- c("Respondents", "Sex", "Fathers Occupation", "Persons At Home", "Siblings A
```

#4. Analyzing the Graph: This bar chart, titled "Sentiment of Tweets per Day," provides a concise overview of the sentiment of tweets posted each day in July 2020.

#Negative Sentiment:

On specific days such as July 15 and July 21, 2020, there was a noticeable increase in negative tweets.

#Neutral Sentiment:

The neutral sentiment represents tweets that maintain an impartial and factual tone. Throughout July 2020, neutral tweets were the most common.

#Positive Sentiment:

Tweets falling into the positive sentiment category are characterized by their enthusiastic and optimistic tone.

#In summary, the bar graph titled "Sentiment of Tweets per Day" provides valuable insights into the emotional tone of tweets posted each day in July 2020.