RWorksheet_cabia#3b

Billy Brendan Cabia

2024-10-02

1.

a.

```
dframe <- 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),
  Fathers_Occupation = c(1, 3, 3, 3, 1, 2, 3, 1, 1, 1, 3, 2, 1, 3, 3, 1, 3, 1, 2, 1),
  Home = 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)
  )

dframe
```

```
##    Respondents Sex Fathers_Occupation Home SiblingsatSchool typesofHouses
## 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(dframe)
```

```
## '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 ...
##  $ Home              : 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 displays the number of rows and columns in the data frame. It provides a preview of the first few entries and shows the data type of each column.

c.

```
mean(dframe$SiblingsatSchool)
```

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

d.

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

```
##   Respondents Sex Fathers_Occupation Home SiblingsatSchool typesofHouses
## 1           1   2                  1    5               6             1
## 2           2   2                  3    7               4             2
```

e.

```
subset <- dframe[c(3, 5), c(2, 4)]
subset
```

```
##   Sex Home
## 3   1    3
## 5   2    5
```

f.

```
types_houses <- dframe$typesofHouses
```

g.

```
maleF <- subset(dframe, Sex == 1 & Fathers_Occupation == 1)
maleF
```

```
## [1] Respondents        Sex                Fathers_Occupation Home
## [5] SiblingsatSchool   typesofHouses
## <0 rows> (or 0-length row.names)
```

h.

```
femaleS <- subset(dframe, Sex == 2 & SiblingsatSchool >= 5)
femaleS
```

```
##    Respondents Sex Fathers_Occupation Home SiblingsatSchool typesofHouses
## 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.

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
```

The code creates an empty data frame with five columns of different data types.

3.

a.

```
householddata <- data.frame(
  Respondents = 1:10,
  Sex = c("Male", "Female", "Female", "Male", "Male", "Female", "Female", "Male", "Female", "Male"),
  Fathers_Occupation = c(1, 2, 3, 3, 1, 2, 2, 1, 1, 3),
  Persons_at_Home = c(5, 7, 3, 8, 6, 4, 2, 4, 11, 6),
  Siblings_at_School = c(5, 3, 3, 5, 6, 3, 1, 2, 6, 6),
  Types_of_Houses = c("Wood", "Concrete", "Concrete", "Wood", "Semi-concrete", "Semi-concrete", "Wood",
)
```

```
householddata
```

```
##    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                  3
## 4            4   Male                  3               8                  5
## 5            5   Male                  1               6                  6
## 6            6 Female                  2               4                  3
## 7            7 Female                  2               2                  1
## 8            8   Male                  1               4                  2
## 9            9 Female                  1              11                  6
## 10          10   Male                  3               6                  6
##    Types_of_Houses
## 1             Wood
## 2         Concrete
## 3         Concrete
## 4             Wood
## 5    Semi-concrete
## 6    Semi-concrete
## 7             Wood
## 8    Semi-concrete
## 9    Semi-concrete
## 10        Concrete
```

```
write.csv(householddata, file = "HouseholdData.csv", row.names = FALSE)

householddata <- read.csv("HouseholdData.csv")

householddata
```

```
##    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                  3
## 4            4   Male                  3               8                  5
## 5            5   Male                  1               6                  6
## 6            6 Female                  2               4                  3
## 7            7 Female                  2               2                  1
## 8            8   Male                  1               4                  2
## 9            9 Female                  1              11                  6
## 10          10   Male                  3               6                  6
##    Types_of_Houses
## 1             Wood
## 2         Concrete
## 3         Concrete
## 4             Wood
## 5    Semi-concrete
## 6    Semi-concrete
## 7             Wood
## 8    Semi-concrete
## 9    Semi-concrete
## 10        Concrete
```

b.

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

householddata$Sex <- as.integer(householddata$Sex)
householddata
```

```
##    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                  3
## 4            4   1                  3               8                  5
## 5            5   1                  1               6                  6
## 6            6   2                  2               4                  3
## 7            7   2                  2               2                  1
## 8            8   1                  1               4                  2
## 9            9   2                  1              11                  6
## 10          10   1                  3               6                  6
##    Types_of_Houses
## 1             Wood
## 2         Concrete
## 3         Concrete
## 4             Wood
## 5    Semi-concrete
## 6    Semi-concrete
## 7             Wood
## 8    Semi-concrete
## 9    Semi-concrete
## 10        Concrete
```

c.

```
householddata$Types_of_Houses <- factor(householddata$Types_of_Houses, levels = c("Wood", "Concrete", "

householddata$Types_of_Houses <- as.integer(householddata$Types_of_Houses)

householddata
```

```
##    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                  3
## 4            4   1                  3               8                  5
## 5            5   1                  1               6                  6
## 6            6   2                  2               4                  3
## 7            7   2                  2               2                  1
## 8            8   1                  1               4                  2
## 9            9   2                  1              11                  6
## 10          10   1                  3               6                  6
##    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$Fathers_Occupation <- as.integer(householddata$Fathers_Occupation)

householddata
```

```
##    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                  3
## 4            4   1                  3               8                  5
## 5            5   1                  1               6                  6
## 6            6   2                  2               4                  3
## 7            7   2                  2               2                  1
## 8            8   1                  1               4                  2
## 9            9   2                  1              11                  6
## 10          10   1                  3               6                  6
##    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_driver_respondents <- subset(householddata, Sex == 2 & Fathers_Occupation == 2)
female_driver_respondents
```

```
##   Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 2           2   2                  2               7                  3
## 6           6   2                  2               4                  3
## 7           7   2                  2               2                  1
##   Types_of_Houses
## 2               2
## 6               3
## 7               1
```

f.

```
siblings <- subset(householddata, Siblings_at_School >= 5)
siblings
```

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

4.

The bar graph shows the sentiment analysis of tweets over a period of days, it shows how the volume of tweets with positive, neutral, and negative sentiments varied daily.