RWorksheet_Francisco#4a

Enrique Francisco

2024-10-14

```
#1.
Household_data <- read.csv("/cloud/project/worksheet#4/Household Data.csv")
Household_data
```

```
##    Shoe.size Height Gender
## 1        6.5   66.0      F
## 2        9.0   68.0      F
## 3        8.5   64.5      F
## 4        8.5   65.0      F
## 5       10.5   70.0      M
## 6        7.0   64.0      F
## 7        9.5   70.0      F
## 8        9.0   71.0      F
## 9       13.0   72.0      M
## 10       7.5   64.0      F
## 11      10.5   74.5      M
## 12       8.5   67.0      F
## 13      12.0   71.0      M
## 14      10.5   71.0      M
## 15      13.0   77.0      M
## 16      11.5   72.0      M
## 17       8.5   59.0      F
## 18       5.0   62.0      F
## 19      10.0   72.0      M
## 20       6.5   66.0      F
## 21       7.5   64.0      F
## 22       8.5   67.0      M
## 23      10.5   73.0      M
## 24       8.5   69.0      F
## 25      10.5   72.0      M
## 26      11.0   70.0      M
## 27       9.0   69.0      M
## 28      13.0   70.0      M
```

#1.a: The data has 28 objects with 3 variables:Shoe size, Height and Gender

```
#1.b
sub1 <- subset(Household_data, Gender == "M" & Shoe.size&Height)
sub1
```

```
##    Shoe.size Height Gender
## 5       10.5   70.0      M
## 9       13.0   72.0      M
## 11      10.5   74.5      M
```

```
## 13      12.0   71.0      M
## 14      10.5   71.0      M
## 15      13.0   77.0      M
## 16      11.5   72.0      M
## 19      10.0   72.0      M
## 22       8.5   67.0      M
## 23      10.5   73.0      M
## 25      10.5   72.0      M
## 26      11.0   70.0      M
## 27       9.0   69.0      M
## 28      13.0   70.0      M
```

```
#1.b
sub2 <- subset(Household_data, Gender == "F" & Shoe.size&Height)
sub2
```

```
##    Shoe.size Height Gender
## 1        6.5   66.0      F
## 2        9.0   68.0      F
## 3        8.5   64.5      F
## 4        8.5   65.0      F
## 6        7.0   64.0      F
## 7        9.5   70.0      F
## 8        9.0   71.0      F
## 10       7.5   64.0      F
## 12       8.5   67.0      F
## 17       8.5   59.0      F
## 18       5.0   62.0      F
## 20       6.5   66.0      F
## 21       7.5   64.0      F
## 24       8.5   69.0      F
```

```
#1.c
mean1 <- mean(Household_data$Shoe.size)
mean1
```

```
## [1] 9.410714
```

```
mean2 <- mean(Household_data$Height)
mean2
```

```
## [1] 68.57143
```

#1.d: Based on the given data if we compare the Male and Female proportion about shoe size and height, there's a big difference for Male because as the Height of Males increase the shoe size also increases while on the other hand Females vary from shoe sizes as there's some who is much shorter than other but have bigger shoe sizes

```
#2
Months <- c("March", "April" ,"January" ,"November" ,"January", "September","October","September","Nover

factor_months_vector <- factor(Months)
factor_months_vector
```

```
##  [1] March     April     January   November  January   September October
##  [8] September November  August    January   November  November  February
## [15] May       August    July      December  August    August    September
## [22] November  February  April
```

```
## 11 Levels: April August December February January July March May ... September
```

```
#3
Sum <- summary(Months)
Sum
```

```
##    Length     Class      Mode
##        24 character character
```

```
Sum2 <- summary(factor_months_vector)
Sum2
```

```
##     April    August  December  February   January      July     March       May
##         2         4         1         2         3         1         1         1
##  November   October September
##         5         1         3
```

```
#4
Datas <- c(c("East", "West", "North"), c(1,4,3))
Datas
```

```
## [1] "East"  "West"  "North" "1"     "4"     "3"
```

```
factor_data <- matrix(Datas,nrow=3,ncol=2)
factor_data
```

```
##      [,1]    [,2]
## [1,] "East"  "1"
## [2,] "West"  "4"
## [3,] "North" "3"
```

```
#4
colnames(factor_data) <- c("Direction", "Frequency")
factor_data
```

```
##      Direction Frequency
## [1,] "East"    "1"
## [2,] "West"    "4"
## [3,] "North"   "3"
```

```
#4
new_order_data <- factor(factor_data,levels = c("East","West","North"))
print(new_order_data)
```

```
## [1] East  West  North <NA>  <NA>  <NA>
## Levels: East West North
```

```
#5a
setwd("/cloud/project/worksheet#4")
Strats <- read.table("import_march.csv", header= TRUE, sep = ",")
Strats
```

```
##   Students Strategy.1 Strategy.2 Strategy.3
## 1     Male          8         10          8
## 2                   4          8          6
## 3                   0          6          4
## 4   Female         14          4         15
## 5                  10          2         12
## 6                   6          0          9
```

```
#5b: The 1,2,3 resulted in NA because those values are missing and did not matched the specified levels
Strats
```

```
##   Students Strategy.1 Strategy.2 Strategy.3
## 1     Male          8         10          8
## 2                   4          8          6
## 3                   0          6          4
## 4   Female         14          4         15
## 5                  10          2         12
## 6                   6          0          9
```