

RWorksheet_Sabarillo#4a

Sabarillo, Kirk Axl Dend

2024-10-14

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

| ## | 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  
male_subset <- subset(household_df, Gender == "M" & Shoe.size&Height)  
male_subset
```

| ## | 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

```
female_subset <- subset(household_df, Gender == "F" & Shoe.size < Height)
female_subset
```

```
##      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_df$Shoe.size)
mean1
```

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

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

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

1.d - Looking at the data, it's clear that the relationship between height and shoe size differs between males and females. For males, there's a more predictable pattern where shoe size increases as height increases. This isn't as consistent for females, where shoe size seems to vary more independently of height.

#2.

```
month_vector <- c("March", "April", "January", "November", "January", "September", "October", "September")
```

```
factor_months <- factor(month_vector)
factor_months
```

```
## [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.
month_summary <- summary(month_vector)
month_summary

##      Length      Class      Mode 
##      24 character character 

factor_month_summary <- summary(factor_months)
factor_month_summary

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

#4.
direction_vector <- c(c("East", "West", "North"), c(1,4,3))
direction_vector

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

direction_matrix <- matrix(direction_vector, nrow=3, ncol=2)
direction_matrix

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

#4.
colnames(direction_matrix) <- c("Direction", "Frequency")
direction_matrix

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

#4.
ordered_direction_factor <- factor(direction_matrix, levels = c("East", "West", "North"))
print(ordered_direction_factor)

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

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

##      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 - Because the numbers 1, 2, and 3 weren't part of the expected options in that category, R marked the output as NA. The correct output is shown below.

```
strat_df
```

```
##   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
```