

# RWorksheet\_Cabia#4a

Billy Brendan Cabia

2024-10-18

1.

```
sframe <- data.frame(  
  Shoe_size = c(6.5, 9.0, 8.5, 8.5, 10.5, 7.0, 9.5, 9.0, 13.0, 7.5, 10.5, 8.5, 12.0, 10.5,  
  Height = c(66.0, 68.0, 64.5, 65.0, 70.0, 64.0, 70.0, 71.0, 72.0, 64.0, 74.5, 67.0, 71.0,  
  Gender = c("F", "F", "F", "F", "M", "F", "F", "F", "M", "F", "M", "F", "M", "M", "M", "M", "M", "F", "F", "  
)  
sframe
```

##	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

a.

The data contains two sets of observations for shoe size, height, and gender labeled as Shoe\_size1, Height1, Gender1 and Shoe\_size, Height, Gender. Each row represents an individual, with the shoe size, height, and gender listed for that person in two separate sets.

b.

```
males <- sframe[sframe$Gender == "M", c("Shoe_size", "Height")]
females <- sframe[sframe$Gender == "F", c("Shoe_size", "Height")]
```

males

##	Shoe_size	Height
## 5	10.5	70.0
## 9	13.0	72.0
## 11	10.5	74.5
## 13	12.0	71.0
## 14	10.5	71.0
## 15	13.0	77.0
## 16	11.5	72.0
## 19	10.0	72.0
## 22	8.5	67.0
## 23	10.5	73.0
## 25	10.5	72.0
## 26	11.0	70.0
## 27	9.0	69.0
## 28	13.0	70.0

females

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

c.

```
mean_shoe_size <- mean(sframe$Shoe_size)
mean_height <- mean(sframe$Height)
```

```
mean_shoe_size
```

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

```
mean_height
```

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

d.

```
correlation <- cor(sframe$Shoe_size, sframe$Height)
correlation
```

```
## [1] 0.7766089
```

2.

```
months_vector <- c(
  "March", "April", "January", "November", "January", "September", "October",
  "September", "November", "August", "January", "November", "November", "February",
  "May", "August", "July", "December", "August", "August", "September", "November",
  "February", "April")
months_vector
```

```
## [1] "March"      "April"      "January"    "November"   "January"    "September"
## [7] "October"    "September"  "November"   "August"     "January"    "November"
## [13] "November"   "February"   "May"        "August"     "July"       "December"
## [19] "August"     "August"     "September"  "November"   "February"   "April"
```

```
factor_months_vector <- factor(months_vector)
```

```
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.

```
summary(months_vector)
```

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

```
summary(factor_months_vector)
```

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

4.

```
directions_vector <- c("East", "West", "North")
frequencies_vector <- c(1, 4, 3)

factor_data <- factor(directions_vector)

new_order_data <- factor(factor_data, levels = c("East", "West", "North"))

new_order_data

## [1] East West North
## Levels: East West North
```