NAME: NAOMI RUTH SALABER BSIT 2A

WORKSHEET 3B

1.A

respondents <- c(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)

persons_at_home <- c(5, 7, 3, 8, 5, 9, 6, 7, 8, 4, 7, 5, 4, 7, 8, 8, 3, 11,7, 6)

siblings_at_school <- c(6, 4, 4, 1, 2, 1, 5, 3, 1, 2, 3, 2, 5, 5, 2, 1, 2, 5, 3, 2)

types_of_houses <- c(1, 2, 3, 1, 1, 3, 3, 1, 2, 3, 2, 3, 2, 2, 3, 3, 3, 3, 3, 2)

r_data <- data.frame(respondents, sex, fathers_occupation,persons_at_home, siblings_at_school,types_of_houses)

r_data

B.

C. No, because the mean of siblings at school is 2.95

D. code: **r_data[1:2, 1:6, drop =FALSE]**

Output:

```
> r_data[1:2, 1:6, drop =FALSE]
  respondents sex fathers_occupation
1           1   2                  1
2           2   2                  3
  persons_at_home siblings_at_school
1               5                  6
2               7                  4
  types_of_houses
1               1
2               2
>
```

E. code**: new_data <- r_data[c(3,5), c(2,4)]**

new_data

output:

```
> new_data <- r_data[c(3,5), c(2,4)]
> new_data
  sex persons_at_home
3   1               3
5   2               5
>
```

F: code: **types_houses <- types_of_houses**

types_houses

output:

```
> types_houses <- types_of_houses
> types_houses
 [1] 1 2 3 1 1 3 3 1 2 3 2 3 2 2 3 3 3 3 3 2
>
```

G. code: **ma**

subset(male

Output:

```
> subset(male_data, sex == 1 & fathers_occupation == 1)
[1] sex                fathers_occupation
<0 rows> (or 0-length row.names)
>
```

H. code: **girl_data <- data.frame(sex, siblings_at_school)**

subset(girl_data, sex == 2 & siblings_at_school >= 5)

Output:

```
> girl_data <- data.frame(sex, siblings_at_school)
> subset(girl_data, sex == 2 & siblings_at_school >= 5)
   sex siblings_at_school
1    2                 6
7    2                 5
13   2                 5
14   2                 5
18   2                 5
>
```

2. A

Answer: the result has 0 obj of 5 variables; the int, num, chr, logi and factors. The result for the levels is NULL.

3. INTERPRET THE GRAPH:

Answer: The graph is about the sentiments of tweets per day and shows that July 15, 2020 has the highest (negative) sentiment from July 14 to 21.