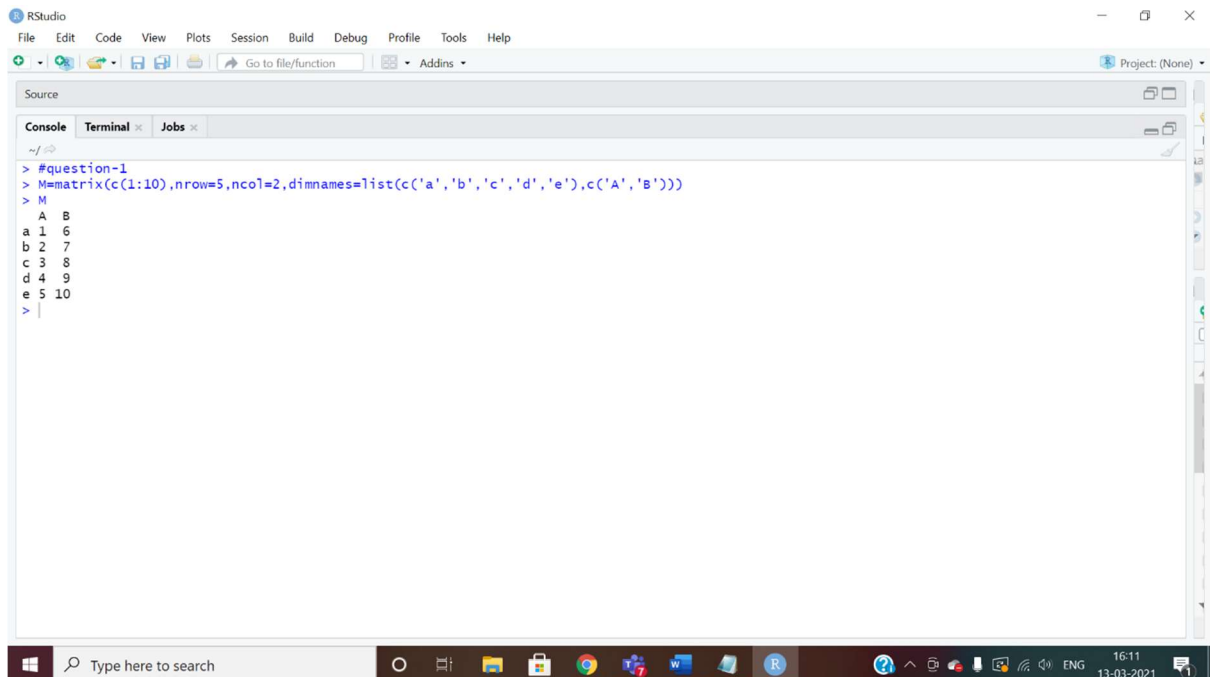


Q1. If `M=matrix(c(1:10),nrow=5,ncol=2, dimnames=list(c('a','b','c','d','e'),c('A','B')))` What is the value of: M

Answer: `M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))`

M

OUTPUT



```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Jobs
~/
> #question-1
> M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B'))
> M
      A B
a 1 6
b 2 7
c 3 8
d 4 9
e 5 10
>
```

Q2. Consider the matrix M, What is the value of: `M[1,]`

`M[,1]`

`M[3,2]`

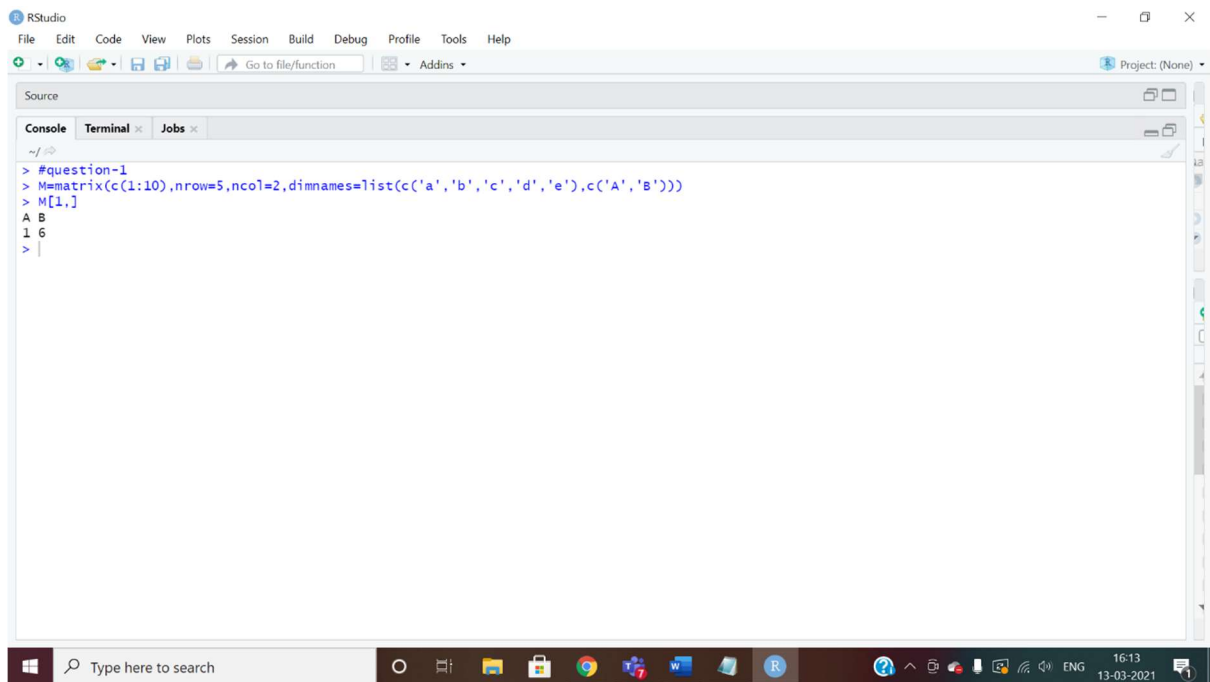
`M['e','A']`

Answer:

1) `M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))`

`M[1,]`

OUTPUT



The image shows the RStudio interface. The console window displays the following code and output:

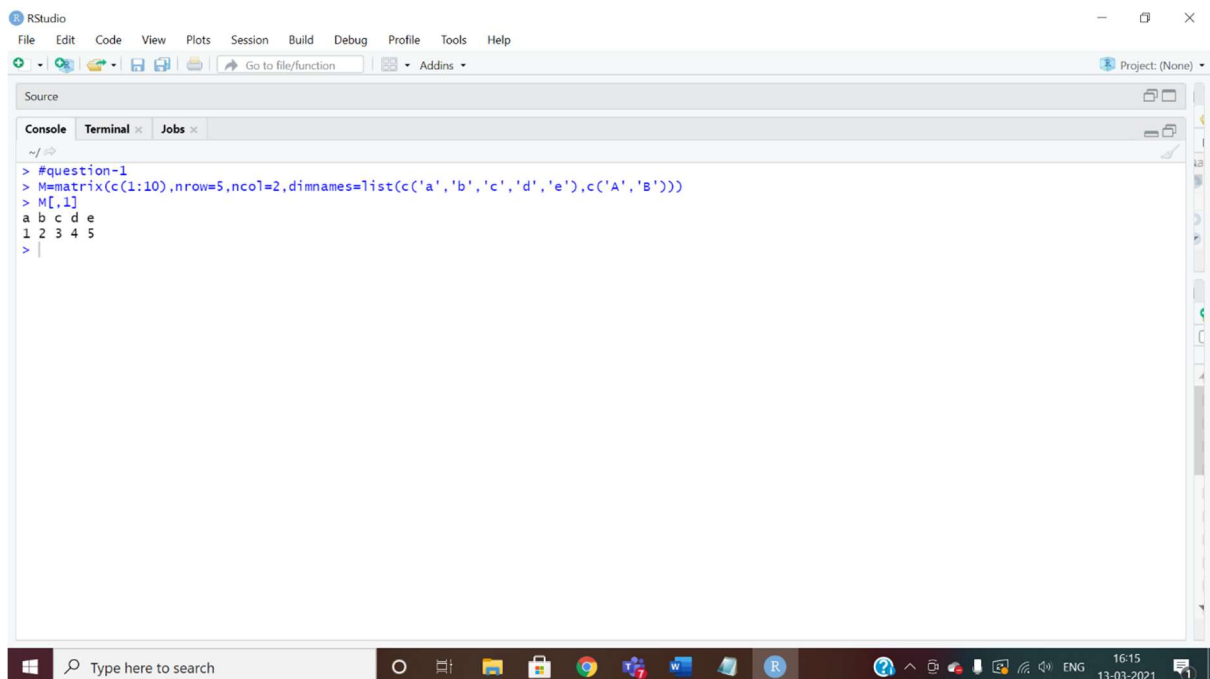
```
> #question-1
> M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))
> M[1,]
A B
1 6
```

The taskbar at the bottom shows the Windows search bar and various application icons. The system clock indicates 16:13 on 13-03-2021.

2) `M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))`

`M[,1]`

OUTPUT



The image shows the RStudio interface. The console window displays the following code and output:

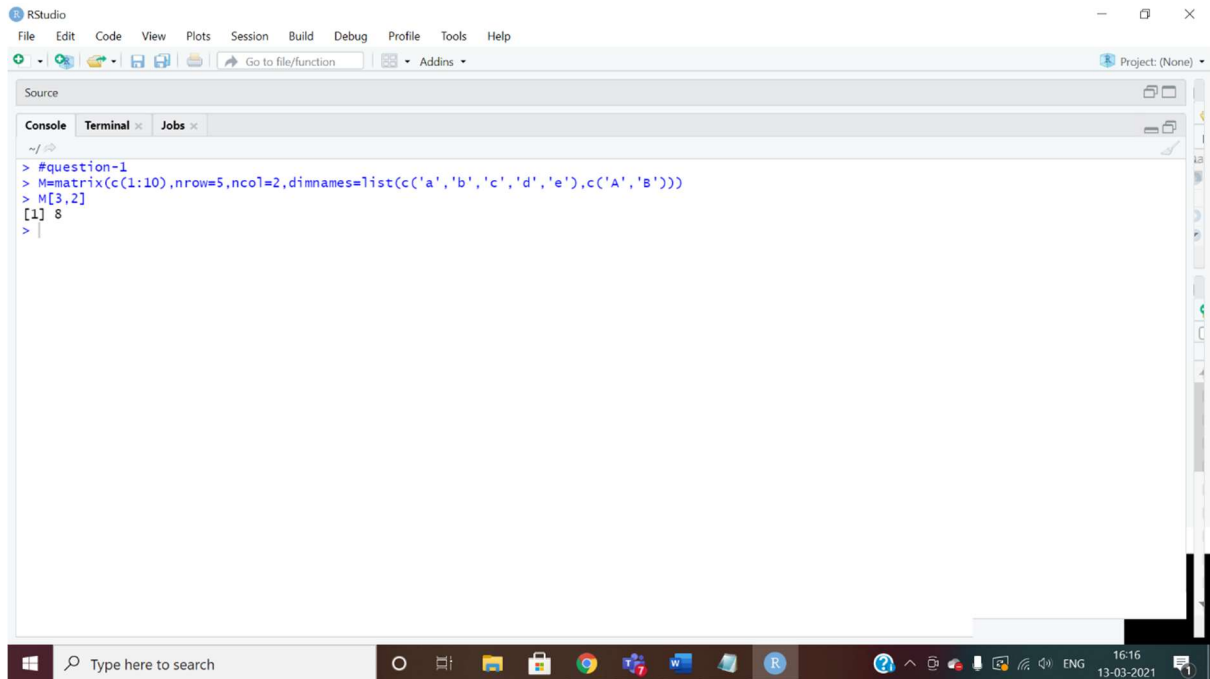
```
> #question-1
> M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))
> M[,1]
a b c d e
1 2 3 4 5
```

The taskbar at the bottom shows the Windows search bar and various application icons. The system clock indicates 16:15 on 13-03-2021.

3) `M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))`

`M[3,2]`

OUTPUT



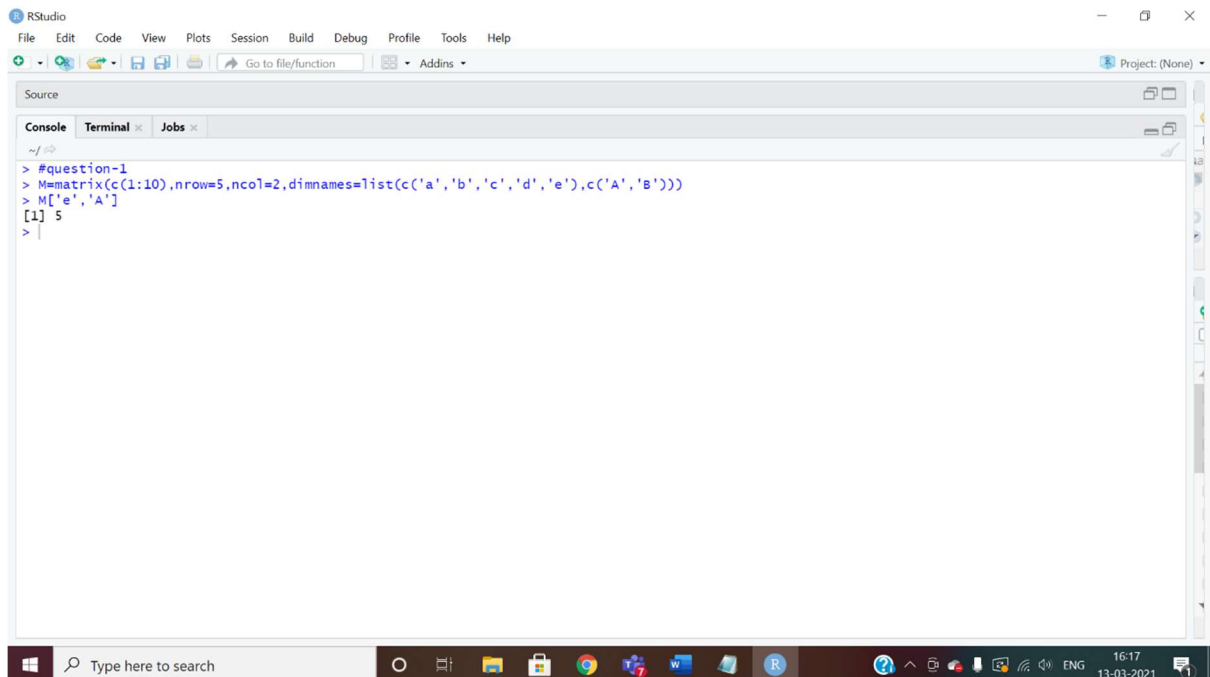
RStudio interface showing the console output for the following R code:

```
> #question-1
> M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))
> M[3,2]
[1] 8
>
```

4) `M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))`

`M['e','A']`

OUTPUT



RStudio interface showing the console output for the following R code:

```
> #question-1
> M=matrix(c(1:10),nrow=5,ncol=2,dimnames=list(c('a','b','c','d','e'),c('A','B')))
> M['e','A']
[1] 5
>
```

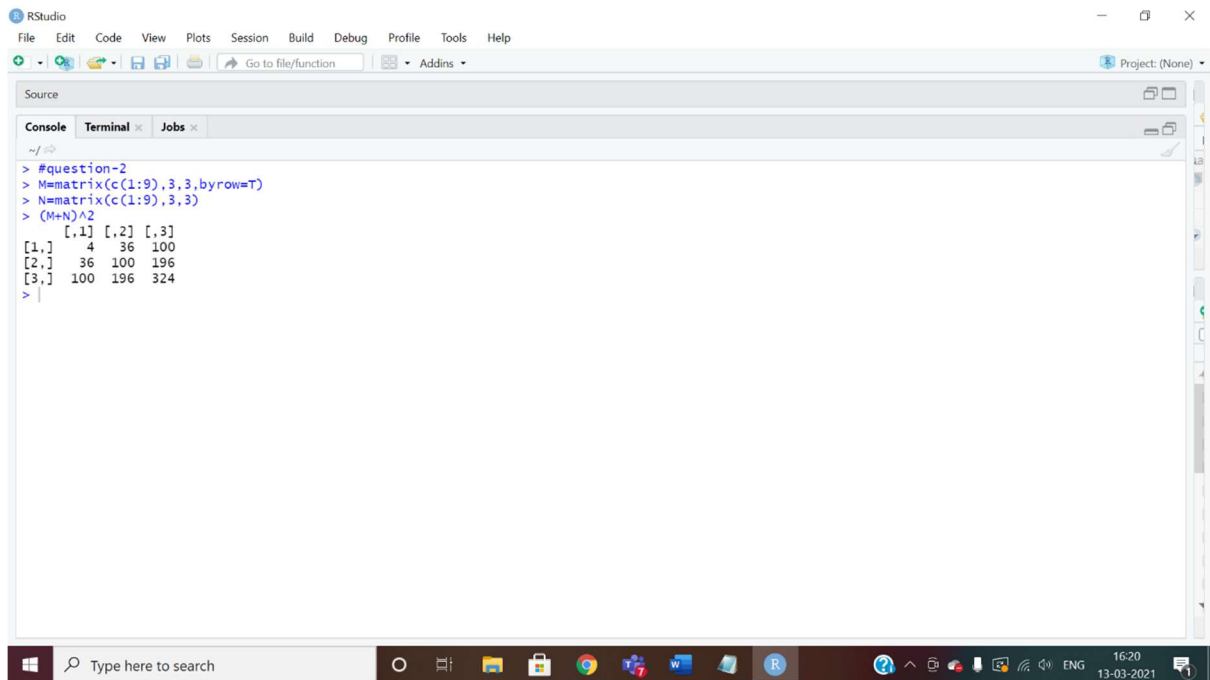
Q2. Consider two matrix, M,N `M=matrix(c(1:9),3,3,byrow=T)` `N=matrix(c(1:9),3,3)` What is the value of: $(M+N)^2$

Answer: `M=matrix(c(1:9),3,3,byrow=T)`

`N=matrix(c(1:9),3,3)`

$(M+N)^2$

OUTPUT



The screenshot shows the RStudio interface with the console pane active. The following R code has been executed:

```
> #question-2  
> M=matrix(c(1:9),3,3,byrow=T)  
> N=matrix(c(1:9),3,3)  
> (M+N)^2
```

The output of the last command is a 3x3 matrix:

```
      [,1] [,2] [,3]  
[1,]    4   36  100  
[2,]   36   64  196  
[3,]  100  196  324
```

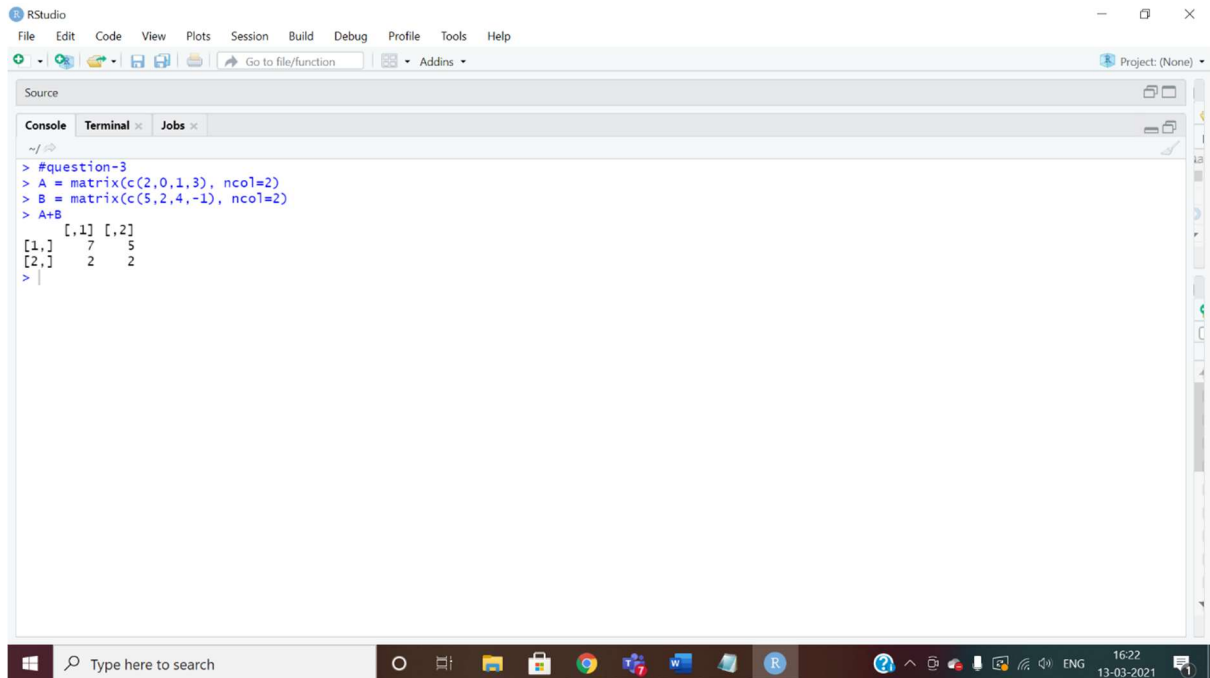
Q3. Consider $A=\text{matrix}(c(2,0,1,3), \text{ncol}=2)$ and $B=\text{matrix}(c(5,2,4,-1), \text{ncol}=2)$. a) Find $A + B$ b) Find $A - B$

Answer:a) $A = \text{matrix}(c(2,0,1,3), \text{ncol}=2)$

$B = \text{matrix}(c(5,2,4,-1), \text{ncol}=2)$

$A+B$

OUTPUT



The screenshot shows the RStudio interface with the console pane active. The code entered is as follows:

```
> #question-3
> A = matrix(c(2,0,1,3), ncol=2)
> B = matrix(c(5,2,4,-1), ncol=2)
> A+B
```

The output displayed in the console is:

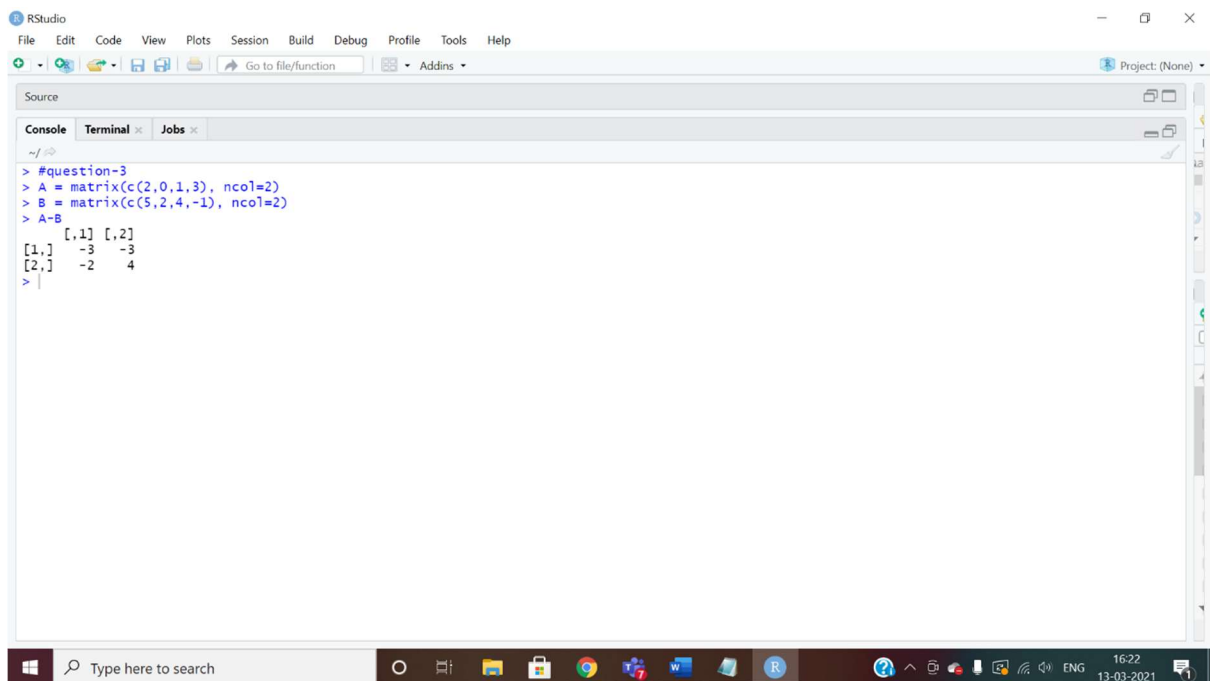
```
      [,1] [,2]
[1,]    7    5
[2,]    2    2
>
```

b) $A = \text{matrix}(c(2,0,1,3), \text{ncol}=2)$

$B = \text{matrix}(c(5,2,4,-1), \text{ncol}=2)$

A-B

OUTPUT



The screenshot shows the RStudio interface with the console pane active. The code entered is as follows:

```
> #question-3
> A = matrix(c(2,0,1,3), ncol=2)
> B = matrix(c(5,2,4,-1), ncol=2)
> A-B
```

The output displayed in the console is:

```
      [,1] [,2]
[1,]   -3   -3
[2,]   -2    4
>
```

Q4. Write the code to create a Data Frame with name df

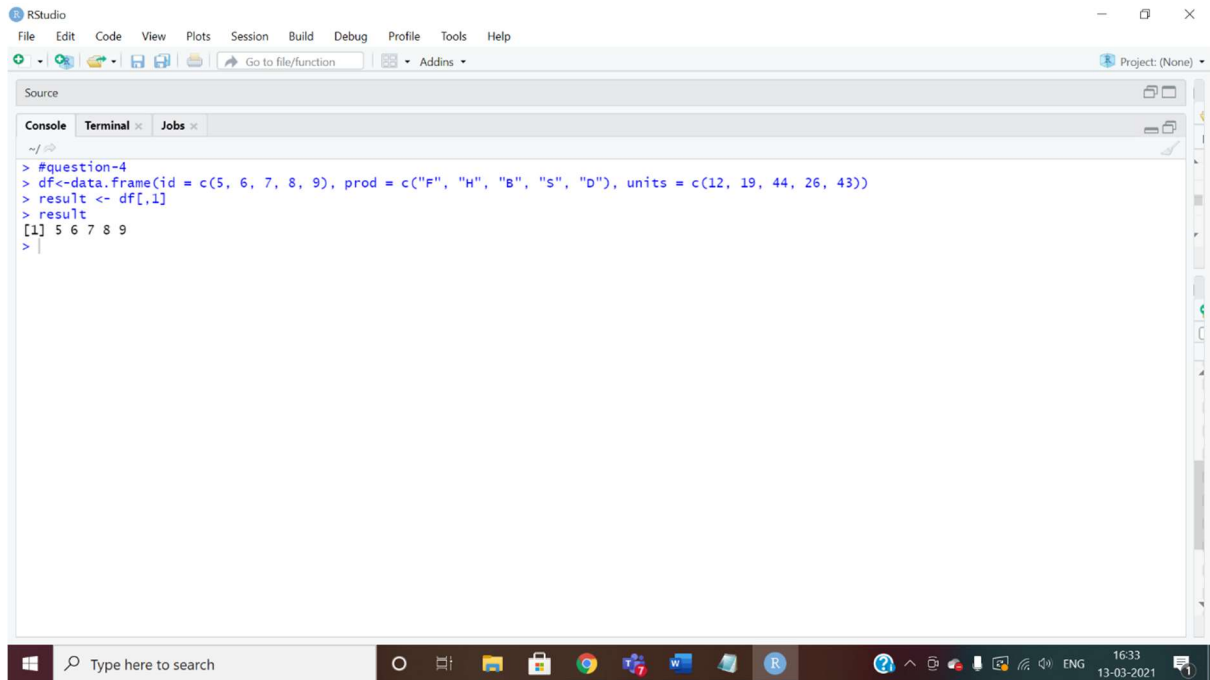
id = c(5, 6, 7, 8, 9), prod = c("F", "H", "B", "S", "D"), units = c(12, 19, 44, 26, 43) How to access the values of attribute id

Answer: `df<-data.frame(id = c(5, 6, 7, 8, 9), prod = c("F", "H", "B", "S", "D"), units = c(12, 19, 44, 26, 43))`

`result <- df[,1]`

`result`

OUTPUT



The screenshot shows the RStudio interface. The console window displays the following R code and its output:

```
> #question-4
> df<-data.frame(id = c(5, 6, 7, 8, 9), prod = c("F", "H", "B", "S", "D"), units = c(12, 19, 44, 26, 43))
> result <- df[,1]
> result
[1] 5 6 7 8 9
> |
```

Q5. Create a dataframe df : `a <- c(rep("A", 3), rep("B", 3), rep("C",2))` `b <- c(1,1,2,4,1,1,2,2)` `df <- data.frame(a,b)`

- Use `duplicated()` function to print the logical vector indicating the duplicate values present in dataframe "df"

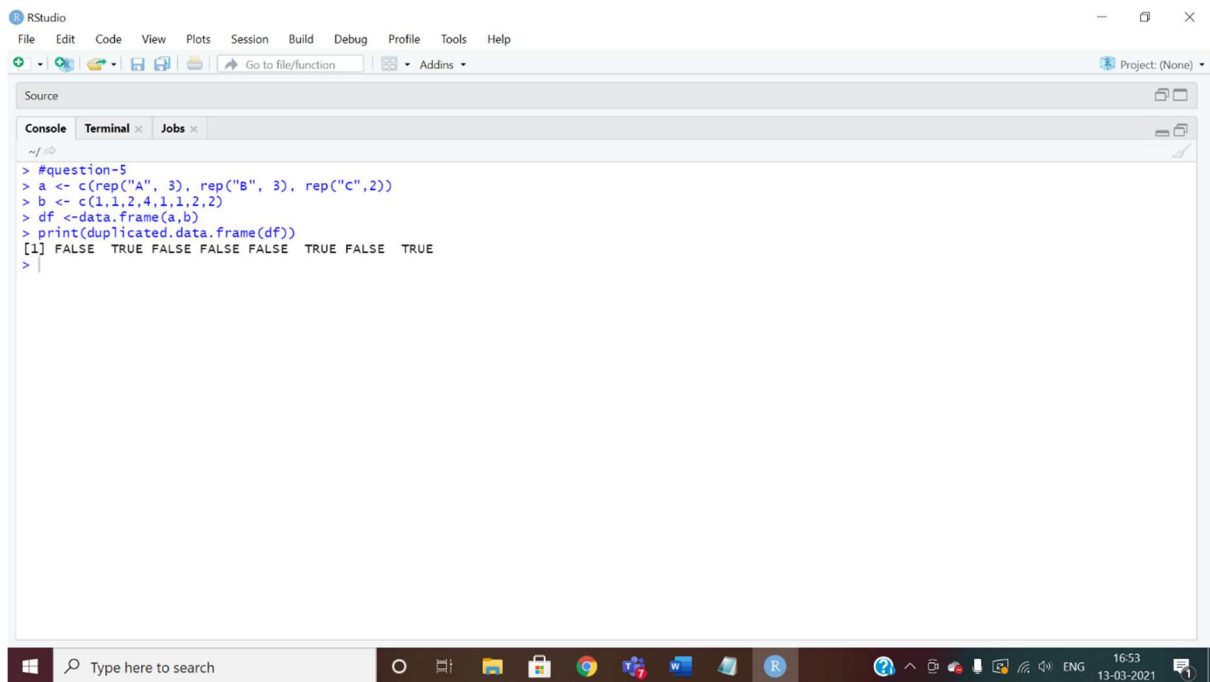
Answer: `a <- c(rep("A", 3), rep("B", 3), rep("C",2))`

`b <- c(1,1,2,4,1,1,2,2)`

`df <-data.frame(a,b)`

`print(duplicated.data.frame(df))`

OUTPUT



The screenshot shows the RStudio interface with the console pane active. The code executed is as follows:

```
#question-5
> a <- c(rep("A", 3), rep("B", 3), rep("C", 2))
> b <- c(1,1,2,4,1,1,2,2)
> df <- data.frame(a,b)
> print(duplicated.data.frame(df))
[1] FALSE TRUE FALSE FALSE FALSE TRUE FALSE TRUE
```

The output indicates that the second and eighth rows of the data frame are duplicated.

- Extract duplicate elements from dataframe "df"

Answer: `a <- c(rep("A", 3), rep("B", 3), rep("C", 2))`

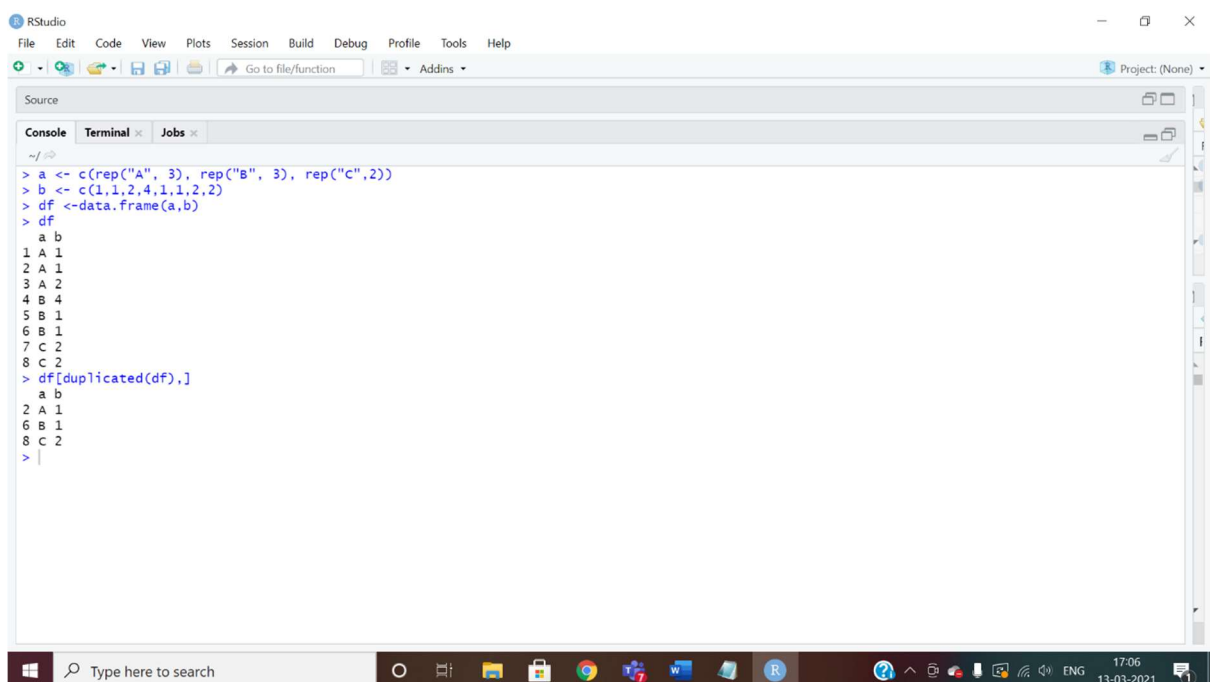
`b <- c(1,1,2,4,1,1,2,2)`

`df <- data.frame(a,b)`

`df`

`df[duplicated(df),]`

OUTPUT



The screenshot shows the RStudio interface with the console pane active. The code executed is as follows:

```
> a <- c(rep("A", 3), rep("B", 3), rep("C", 2))
> b <- c(1,1,2,4,1,1,2,2)
> df <- data.frame(a,b)
> df
  a b
1 A 1
2 A 1
3 A 2
4 B 4
5 B 1
6 B 1
7 C 2
8 C 2
> df[duplicated(df),]
  a b
2 A 1
6 B 1
8 C 2
```

The output shows the original data frame 'df' with 8 rows and the subset of duplicated rows, which are rows 2, 6, and 8.

- Extract unique elements from dataframe "df"

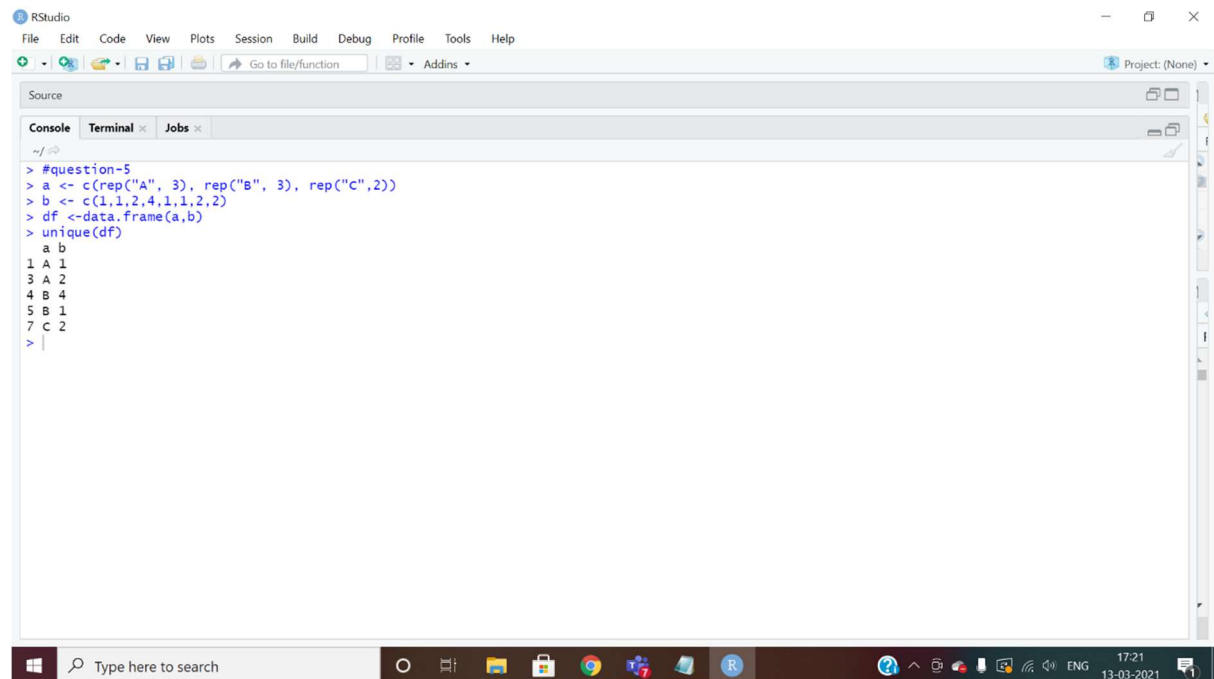
Answer: `a <- c(rep("A", 3), rep("B", 3), rep("C", 2))`

`b <- c(1,1,2,4,1,1,2,2)`

`df <- data.frame(a,b)`

`unique(df)`

OUTPUT



The screenshot shows the RStudio interface with the console window open. The following R code has been executed:

```
> #question-5
> a <- c(rep("A", 3), rep("B", 3), rep("C", 2))
> b <- c(1,1,2,4,1,1,2,2)
> df <- data.frame(a,b)
> unique(df)
```

The output of the `unique(df)` command is displayed in the console:

```
  a b
1 A 1
3 A 2
4 B 4
5 B 1
7 C 2
```

- Print the indices of duplicate elements

Answer: `a <- c(rep("A",3),rep("B",3),rep("C",2))`

`b <- c(1,1,2,4,1,1,2,2)`

`df<-data.frame(a,b)`

`which(duplicated(df))`

OUTPUT


```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Jobs
~/
> a <- c(rep("A",3),rep("B",3),rep("C",2))
> b <- c(1,1,2,4,1,1,2,2)
> df<-data.frame(a,b)
> which(duplicated(df))
[1] 2 6 8
>
```

- Print the indices of unique elements

Answer: a <- c(rep("A",3),rep("B",3),rep("C",2))

b <- c(1,1,2,4,1,1,2,2)

df<-data.frame(a,b)

which(!duplicated(df))

OUTPUT

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Jobs
~/
> #question-5
> a <- c(rep("A",3),rep("B",3),rep("C",2))
> b <- c(1,1,2,4,1,1,2,2)
> df<-data.frame(a,b)
> which(!duplicated(df))
[1] 1 3 4 5 7
>
```

- How many unique elements are in dataframe "df"

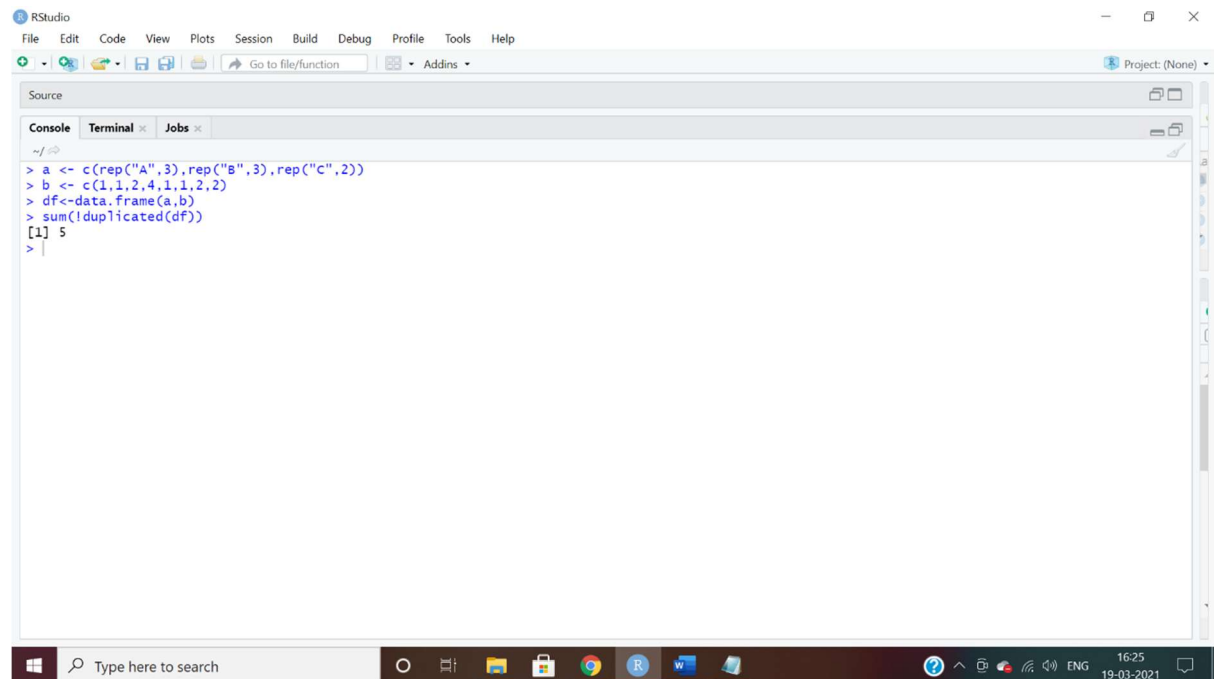
Answer: a <- c(rep("A",3),rep("B",3),rep("C",2))

b <- c(1,1,2,4,1,1,2,2)

df<-data.frame(a,b)

sum(!duplicated(df))

OUTPUT



The screenshot shows the RStudio interface with the console pane active. The following R code has been executed:

```
> a <- c(rep("A",3),rep("B",3),rep("C",2))
> b <- c(1,1,2,4,1,1,2,2)
> df<-data.frame(a,b)
> sum(!duplicated(df))
[1] 5
>
```

The output of the last command is 5, indicating that there are 5 unique elements in the dataframe 'df'.

- How many duplicate elements are in dataframe "df"

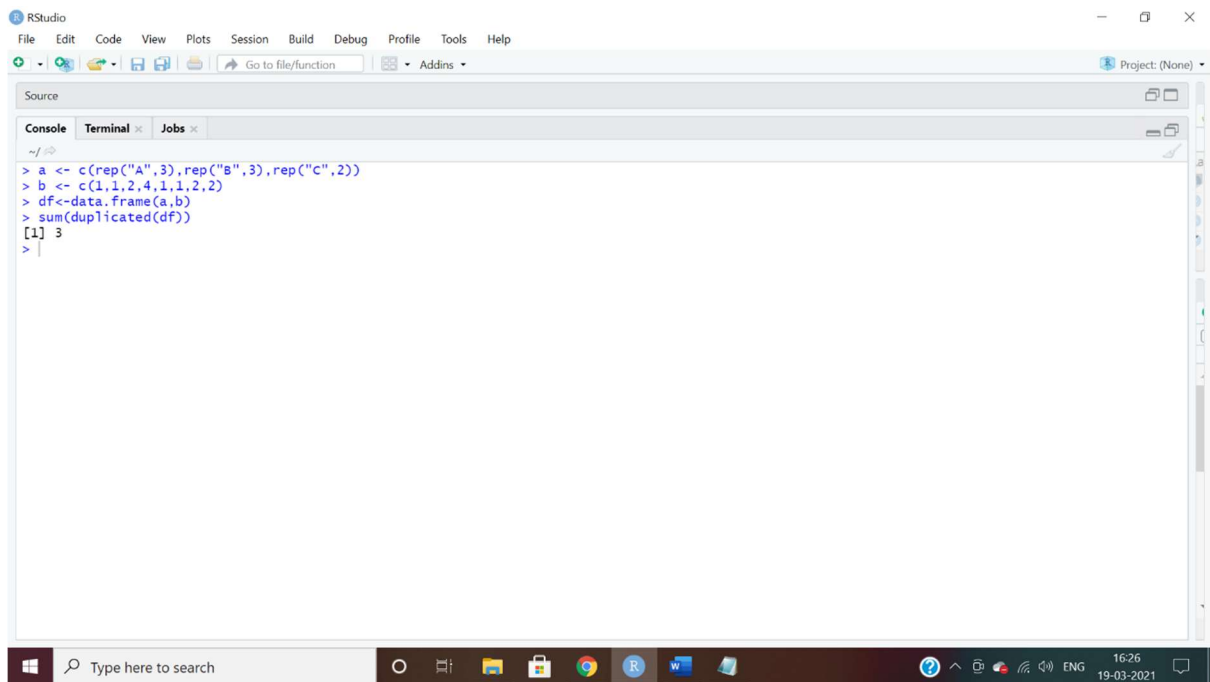
Answer: a <- c(rep("A",3),rep("B",3),rep("C",2))

b <- c(1,1,2,4,1,1,2,2)

df<-data.frame(a,b)

sum(duplicated(df))

OUTPUT

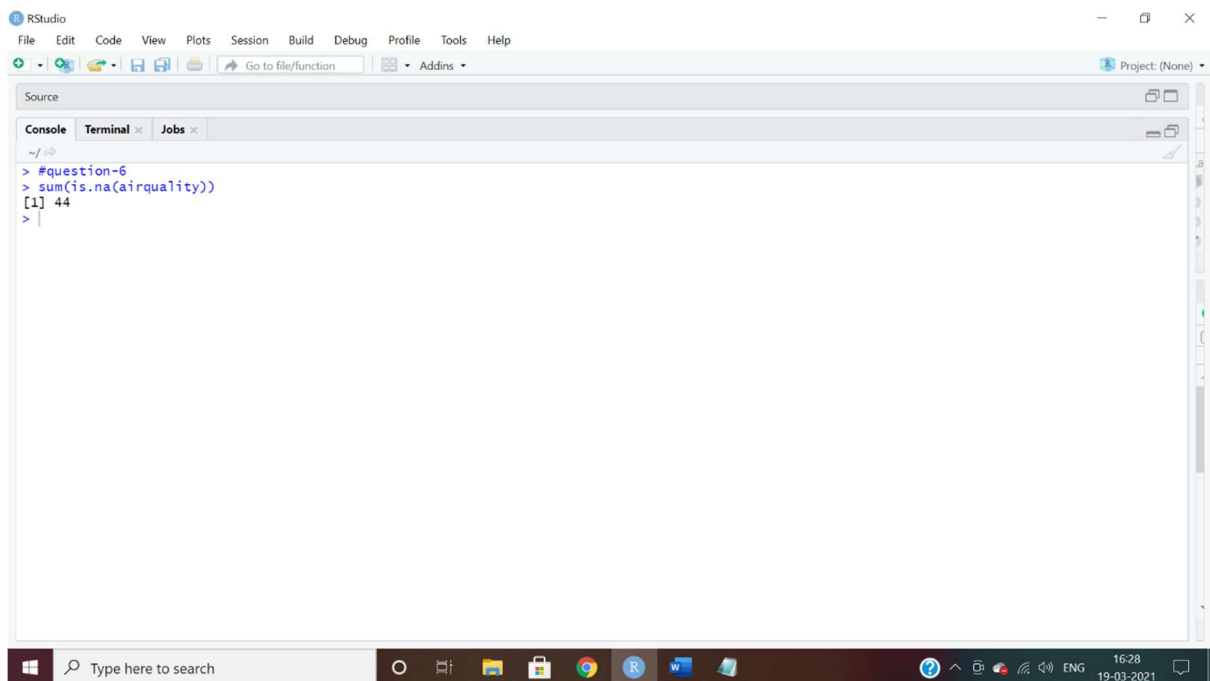


6. Print the dataset airquality

- How many missing values are in the dataset airquality?

Answer: `sum(is.na(airquality))`

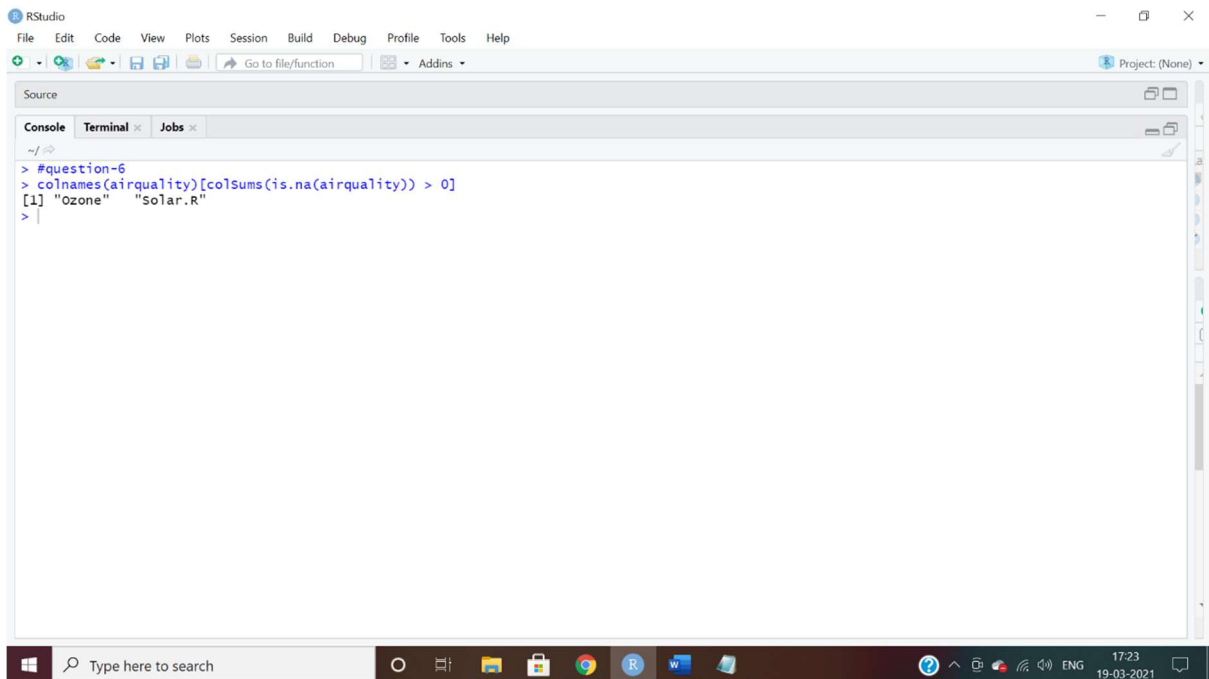
OUTPUT



- Which variables are the missing values concentrated in?

Answer: `colnames(airquality)[colSums(is.na(airquality)) > 0]`

OUTPUT

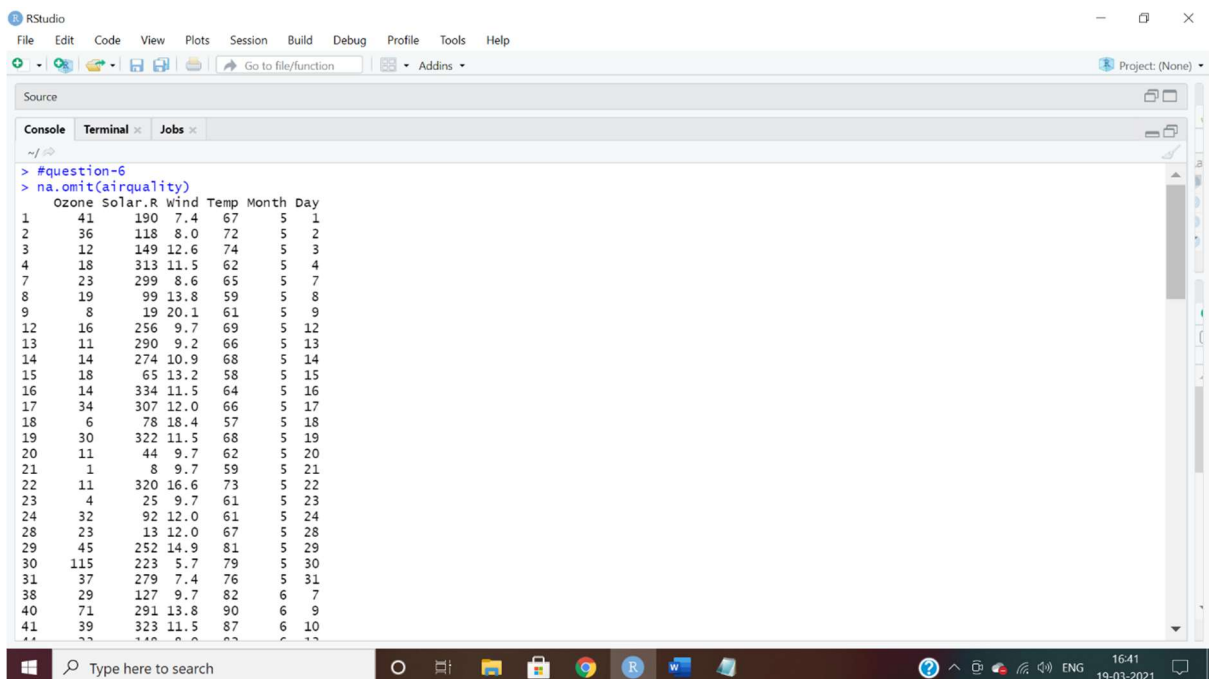


```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Jobs
~/
> #question-6
> colnames(airquality)[colSums(is.na(airquality)) > 0]
[1] "ozone" "solar.R"
>
```

- How would you omit all rows containing missing values?

Answer: `na.omit(airquality)`

OUTPUT

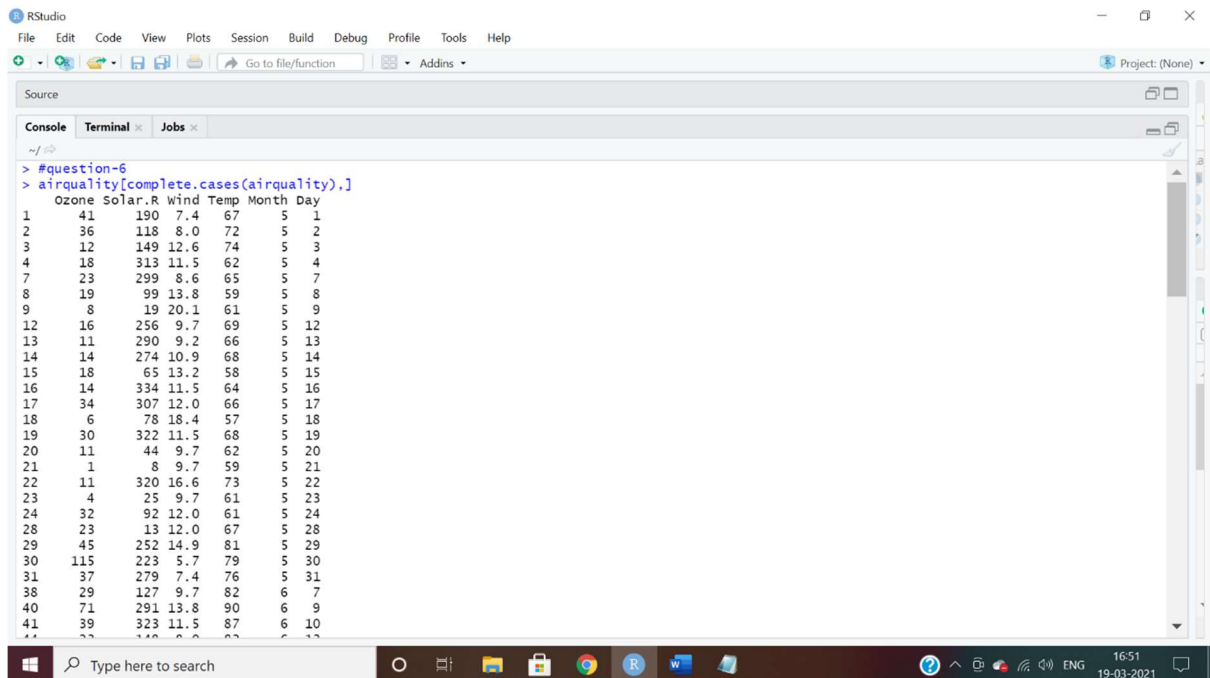


```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Jobs
~/
> #question-6
> na.omit(airquality)
  Ozone solar.R Wind Temp Month Day
1    41    190  7.4   67     5    1
2    36    118  8.0   72     5    2
3    12    149 12.6   74     5    3
4    18    313 11.5   62     5    4
7    23    299  8.6   65     5    7
8    19     99 13.8   59     5    8
9     8     19 20.1   61     5    9
12   16    256  9.7   69     5   12
13   11    290  9.2   66     5   13
14   14    274 10.9   68     5   14
15   18     65 13.2   58     5   15
16   14    334 11.5   64     5   16
17   34    307 12.0   66     5   17
18    6     78 18.4   57     5   18
19   30    322 11.5   68     5   19
20   11     44  9.7   62     5   20
21    1      8  9.7   59     5   21
22   11    320 16.6   73     5   22
23    4     25  9.7   61     5   23
24   32     92 12.0   61     5   24
28   23     13 12.0   67     5   28
29   45    252 14.9   81     5   29
30  115    223  5.7   79     5   30
31   37    279  7.4   76     5   31
38   29    127  9.7   82     6    7
40   71    291 13.8   90     6    9
41   39    323 11.5   87     6   10
```

- Print the records without missing values in the dataset `airquality` using the function `complete.cases()`

Answer: `airquality[complete.cases(airquality),]`

OUTPUT



The screenshot shows the RStudio interface with the Console pane active. The following R code has been executed:

```
> #question-6
> airquality[complete.cases(airquality),]
```

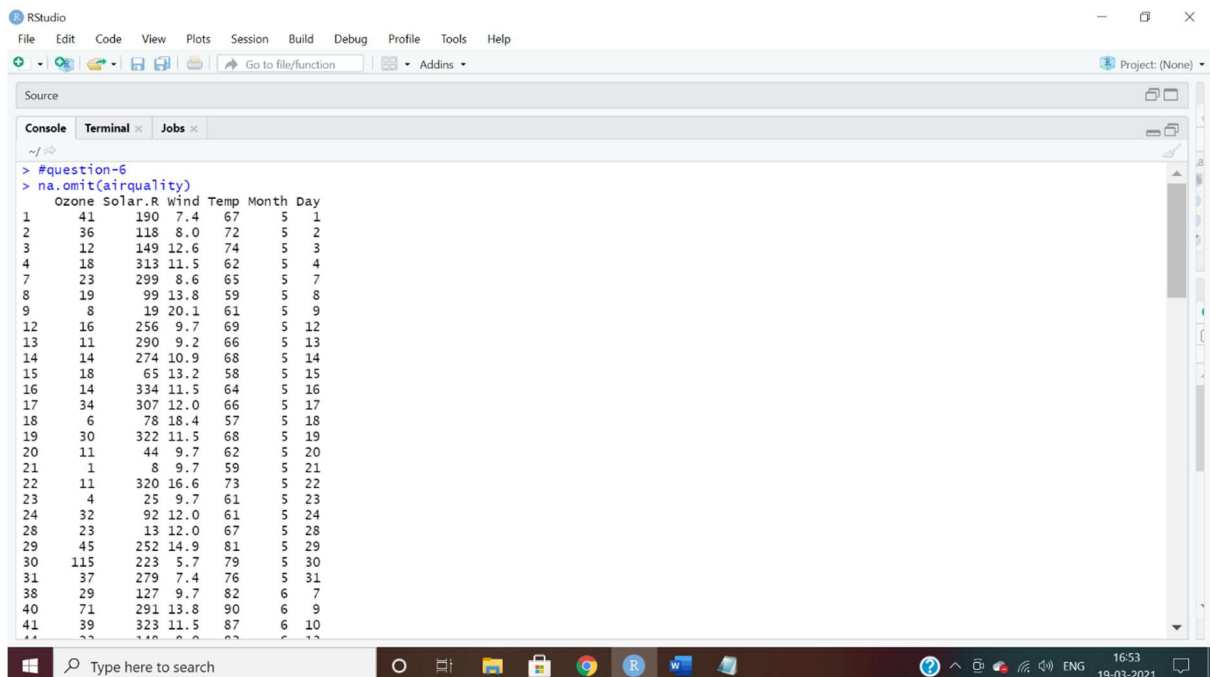
The output is a data frame with 41 rows and 6 columns: Ozone, solar.R, Wind, Temp, Month, and Day. The data represents the first 41 non-missing observations from the 'airquality' dataset.

	Ozone	solar.R	Wind	Temp	Month	Day
1	41	190	7.4	67	5	1
2	36	118	8.0	72	5	2
3	12	149	12.6	74	5	3
4	18	313	11.5	62	5	4
7	23	299	8.6	65	5	7
8	19	99	13.8	59	5	8
9	8	19	20.1	61	5	9
12	16	256	9.7	69	5	12
13	11	290	9.2	66	5	13
14	14	274	10.9	68	5	14
15	18	65	13.2	58	5	15
16	14	334	11.5	64	5	16
17	34	307	12.0	66	5	17
18	6	78	18.4	57	5	18
19	30	322	11.5	68	5	19
20	11	44	9.7	62	5	20
21	1	8	9.7	59	5	21
22	11	320	16.6	73	5	22
23	4	25	9.7	61	5	23
24	32	92	12.0	61	5	24
28	23	13	12.0	67	5	28
29	45	252	14.9	81	5	29
30	115	223	5.7	79	5	30
31	37	279	7.4	76	5	31
38	29	127	9.7	82	6	7
40	71	291	13.8	90	6	9
41	39	323	11.5	87	6	10

- Print the records without missing values in the dataset `airquality` using the function `na.omit()`

Answer: `na.omit(airquality)`

OUTPUT



The screenshot shows the RStudio interface with the Console pane active. The following R code has been executed:

```
> #question-6
> na.omit(airquality)
```

The output is a data frame with 41 rows and 6 columns: Ozone, solar.R, Wind, Temp, Month, and Day. This output is identical to the one in the previous screenshot, as the 'airquality' dataset does not contain any missing values.

	Ozone	solar.R	Wind	Temp	Month	Day
1	41	190	7.4	67	5	1
2	36	118	8.0	72	5	2
3	12	149	12.6	74	5	3
4	18	313	11.5	62	5	4
7	23	299	8.6	65	5	7
8	19	99	13.8	59	5	8
9	8	19	20.1	61	5	9
12	16	256	9.7	69	5	12
13	11	290	9.2	66	5	13
14	14	274	10.9	68	5	14
15	18	65	13.2	58	5	15
16	14	334	11.5	64	5	16
17	34	307	12.0	66	5	17
18	6	78	18.4	57	5	18
19	30	322	11.5	68	5	19
20	11	44	9.7	62	5	20
21	1	8	9.7	59	5	21
22	11	320	16.6	73	5	22
23	4	25	9.7	61	5	23
24	32	92	12.0	61	5	24
28	23	13	12.0	67	5	28
29	45	252	14.9	81	5	29
30	115	223	5.7	79	5	30
31	37	279	7.4	76	5	31
38	29	127	9.7	82	6	7
40	71	291	13.8	90	6	9
41	39	323	11.5	87	6	10

- Print the records containing missing values in the dataset `airquality` using the function `complete.cases()`

Answer: `airquality[complete.cases(airquality)| !complete.cases(airquality),]`

OUTPUT

The screenshot shows the RStudio interface with the Console pane active. The following R code has been executed:

```
> #question-6
> airquality[complete.cases(airquality)| !complete.cases(airquality),]
```

The output displays a subset of the `airquality` dataset, showing rows where any variable has a missing value (NA). The columns are `Ozone`, `Solar.R`, `Wind`, `Temp`, `Month`, and `Day`. The rows are numbered 1 through 27.

	Ozone	Solar.R	Wind	Temp	Month	Day
1	41	190	7.4	67	5	1
2	36	118	8.0	72	5	2
3	12	149	12.6	74	5	3
4	18	313	11.5	62	5	4
5	NA	NA	14.3	56	5	5
6	28	NA	14.9	66	5	6
7	23	299	8.6	65	5	7
8	19	99	13.8	59	5	8
9	8	19	20.1	61	5	9
10	NA	194	8.6	69	5	10
11	7	NA	6.9	74	5	11
12	16	256	9.7	69	5	12
13	11	290	9.2	66	5	13
14	14	274	10.9	68	5	14
15	18	65	13.2	58	5	15
16	14	334	11.5	64	5	16
17	34	307	12.0	66	5	17
18	6	78	18.4	57	5	18
19	30	322	11.5	68	5	19
20	11	44	9.7	62	5	20
21	1	8	9.7	59	5	21
22	11	320	16.6	73	5	22
23	4	25	9.7	61	5	23
24	32	92	12.0	61	5	24
25	NA	66	16.6	57	5	25
26	NA	266	14.9	58	5	26
27	NA	NA	8.0	57	5	27

7. Consider the given vectors: `A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)` `B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)`

- Find the length of the vector A

Answer: `A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)`

`length(A)`

OUTPUT

The screenshot shows the RStudio interface. The console pane displays the following R code and its output:

```
> A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)
> length(A)
[1] 11
> |
```

The top of the window shows the RStudio menu bar (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help) and a toolbar. The bottom status bar shows the Windows taskbar with a search bar and system icons, including the date and time (17:28, 19-03-2021).

- Find the length of the vector B

Answer: B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)

length(B)

OUTPUT

The screenshot shows the RStudio interface. The console pane displays the following R code and its output:

```
> #question-7
> B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)
> length(B)
[1] 11
> |
```

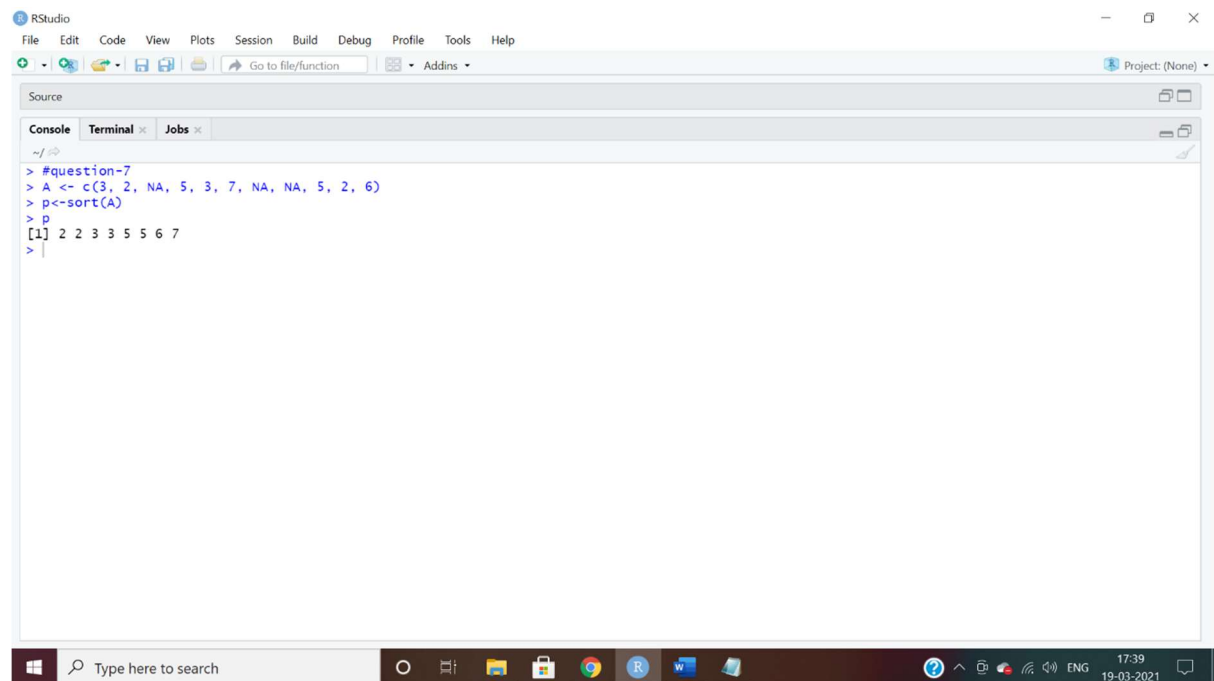
The top of the window shows the RStudio menu bar (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help) and a toolbar. The bottom status bar shows the Windows taskbar with a search bar and system icons, including the date and time (17:38, 19-03-2021).

- Sort the values in vector A and put it in p

```
A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)
```

```
p<-sort(A)
```

OUTPUT



The screenshot shows the RStudio interface with the console pane active. The following R code has been executed:

```
> #question-7  
> A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)  
> p<-sort(A)  
> p  
[1] 2 2 3 3 5 5 6 7  
>
```

The output shows the sorted values of vector A, with NA values removed. The console window is titled 'Console' and has tabs for 'Terminal' and 'Jobs'. The RStudio menu bar and toolbar are visible at the top, and the Windows taskbar is at the bottom.

- Find the length of p

```
Answer: A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)
```

```
p<-sort(A)
```

```
length(p)
```

OUTPUT

The image shows the RStudio interface with the console pane active. The code entered is as follows:

```
> #question-7
> A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)
> p<-sort(A)
> length(p)
[1] 8
>
```

The Windows taskbar at the bottom shows the time as 17:39 on 19-03-2021.

- Sort the values in vector B and put it in q

Answer: B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)

q<-sort(B)

OUTPUT

The image shows the RStudio interface with the console pane active. The code and its output are as follows:

```
> #question-7
> B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)
> q<-sort(B)
> q
[1] "2" "2" "3" "3" "5" "5" "6" "7" "NA"
>
```

The Windows taskbar at the bottom shows the time as 17:41 on 19-03-2021.

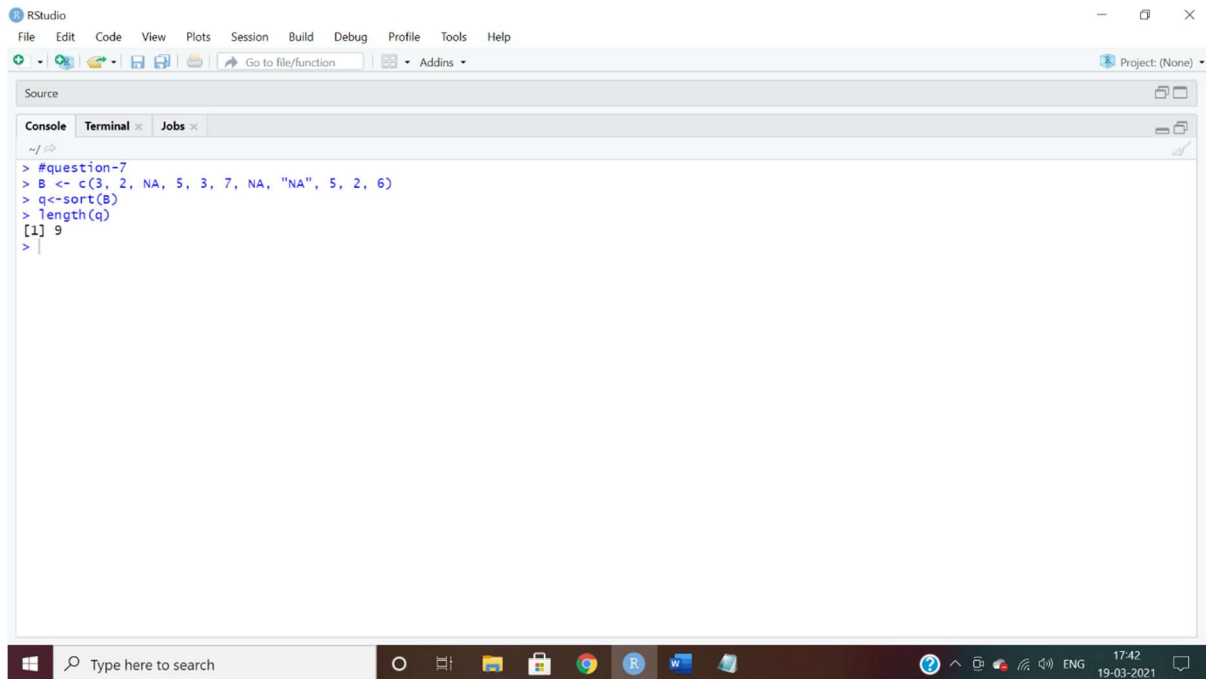
- Find the length of q

Answer: `B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)`

`q<-sort(B)`

`length(q)`

OUTPUT



The screenshot shows the RStudio interface with the console pane active. The following R code was executed:

```
> #question-7  
> B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)  
> q<-sort(B)  
> length(q)  
[1] 9  
> |
```

The output of the `length(q)` command is `[1] 9`. The RStudio window title is "RStudio" and the project is "(None)". The Windows taskbar at the bottom shows the date and time as 17:42 on 19-03-2021.

- What did you infer from the above results

Answer: **#Inferring**

#Sort function removes NA, but since in B, One NA is taken as character, it's treated as character and hence the length is 1 more than A.