Problem Statement

1.

df1 = data.frame(CustId = c(1:6), Product = c(rep("TV", 3), rep("Radio", 3)))
df2 = data.frame(CustId = c(2, 4, 6), State = c(rep("Texas", 2), rep("NYC", 1)))
df1 #left table
df2 #right table

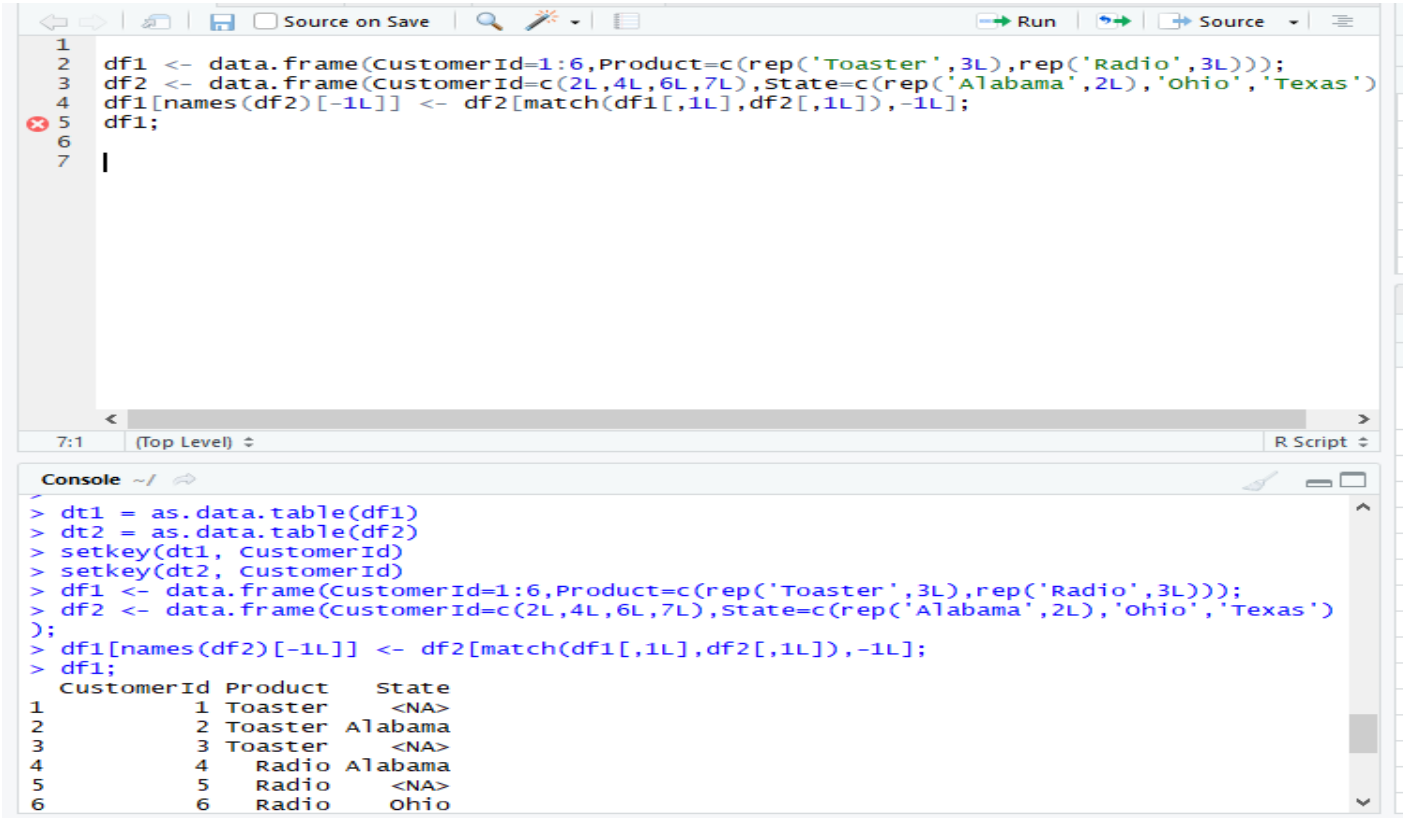
For the above given data frames and tables perform the following operations:

- **Return only the rows in which the left table have match.**
- **Returns all rows from both tables, join records from the left which have matching keys in the right table.**
- **Return all rows from the left table, and any rows with matching keys from the right table.**
- **Return all rows from the right table, and any rows with matching keys from the left table.**

Answer :

```
df1 <- data.frame(CustomerId=1:6,Product=c(rep('Toaster',3L),rep('Radio',3L)));
df2 <- data.frame(CustomerId=c(2L,4L,6L,7L),State=c(rep('Alabama',2L),'Ohio','Texas'));
df1[names(df2)[-1L]] <- df2[match(df1[,1L],df2[,1L]),-1L];
df1;
df1;
  CustomerId Product State
1          1 Toaster  <NA>
2          2 Toaster Alabama
3          3 Toaster  <NA>
4          4 Radio Alabama
5          5 Radio    <NA>
6          6 Radio Ohio

> df2;
   CustomerId State
1           2 Alabama
2           4 Alabama
3           6 Ohio
4           7 Texas
```



merge(df1, df2)
CustomerId State Product
1 2 Alabama Toaster
2 4 Alabama Radio
3 6 Ohio Radio
> merge(df1, df2, by = "CustomerId")
CustomerId Product State.x State.y
1 2 Toaster Alabama Alabama
2 4 Radio Alabama Alabama
3 6 Radio Ohio Ohio

2. **Perform the below operations on above given data frames and tables:**

- **Return a long format of the datasets without matching key.**
- **Keep only observations in df1 that match in df2.**
- **Drop all observations in df1 that match in df2.**

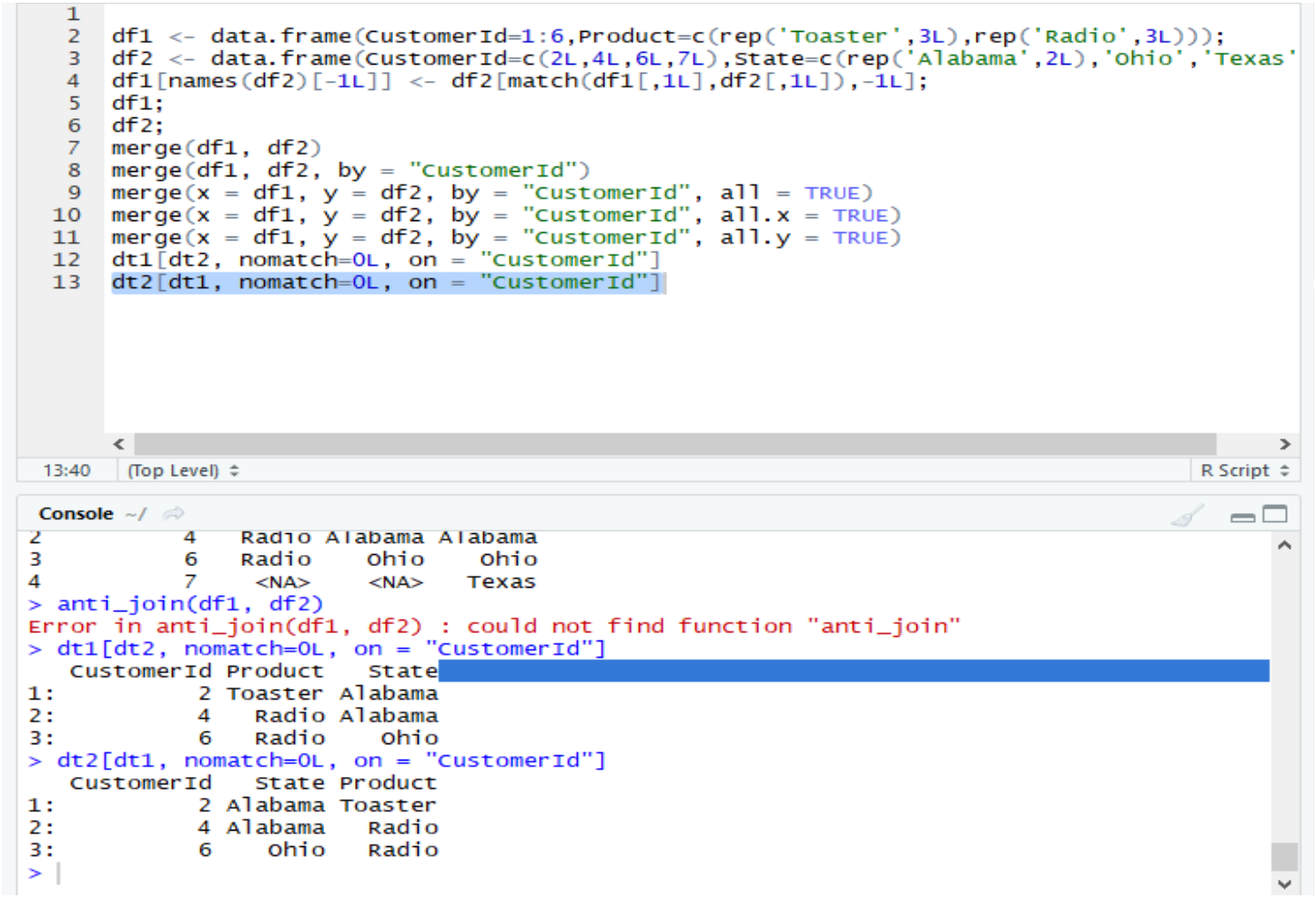
Answer :

dt1[dt2, nomatch=0L, on = "CustomerId"]

> dt1[dt2, nomatch=0L, on = "CustomerId"]
CustomerId Product State
1: 2 Toaster Alabama
2: 4 Radio Alabama

3: 6 Radio Ohio

dt2[dt1, nomatch=0L, on = "CustomerId"]
> dt2[dt1, nomatch=0L, on = "CustomerId"]
CustomerId State Product
1: 2 Alabama Toaster
2: 4 Alabama Radio
3: 6 Ohio Radio



```
> dt1[dt2, nomatch=0L, on = "CustomerId"]
   CustomerId Product   State
1:          2 Toaster Alabama
2:          4   Radio Alabama
3:          6   Radio    Ohio
> dt2[dt1, nomatch=0L, on = "CustomerId"]
   CustomerId   State Product
1:          2 Alabama Toaster
2:          4 Alabama   Radio
3:          6    Ohio   Radio
> merge(x = df1, y = df2, by = NULL)
   CustomerId.x Product State.x CustomerId.y State.y
1             1 Toaster    <NA>            2 Alabama
2             2 Toaster Alabama            2 Alabama
3             3 Toaster    <NA>            2 Alabama
4             4   Radio Alabama            2 Alabama
5             5   Radio    <NA>            2 Alabama
6             6   Radio    Ohio            2 Alabama
7             1 Toaster    <NA>            4 Alabama
8             2 Toaster Alabama            4 Alabama
9             3 Toaster    <NA>            4 Alabama
10            4   Radio Alabama            4 Alabama
11            5   Radio    <NA>            4 Alabama
12            6   Radio    Ohio            4 Alabama
13            1 Toaster    <NA>            6    Ohio
14            2 Toaster Alabama            6    Ohio
15            3 Toaster    <NA>            6    Ohio
16            4   Radio Alabama            6    Ohio
17            5   Radio    <NA>            6    Ohio
18            6   Radio    Ohio            6    Ohio
19            1 Toaster    <NA>            7   Texas
20            2 Toaster Alabama            7   Texas
21            3 Toaster    <NA>            7   Texas
22            4   Radio Alabama            7   Texas
23            5   Radio    <NA>            7   Texas
24            6   Radio    Ohio            7   Texas
>
```