ACADGILD

SESSION 4: FOUNDATIONAL R PROGRAMMING-II

Assignment 1

Table of Contents

1. Introduction

This assignment will help you understand the concepts learnt in the session.

2. Objective

This assignment will test your skills on foundational R Programming- writing functions.

3. Prerequisites

Not applicable.

4. Associated Data Files
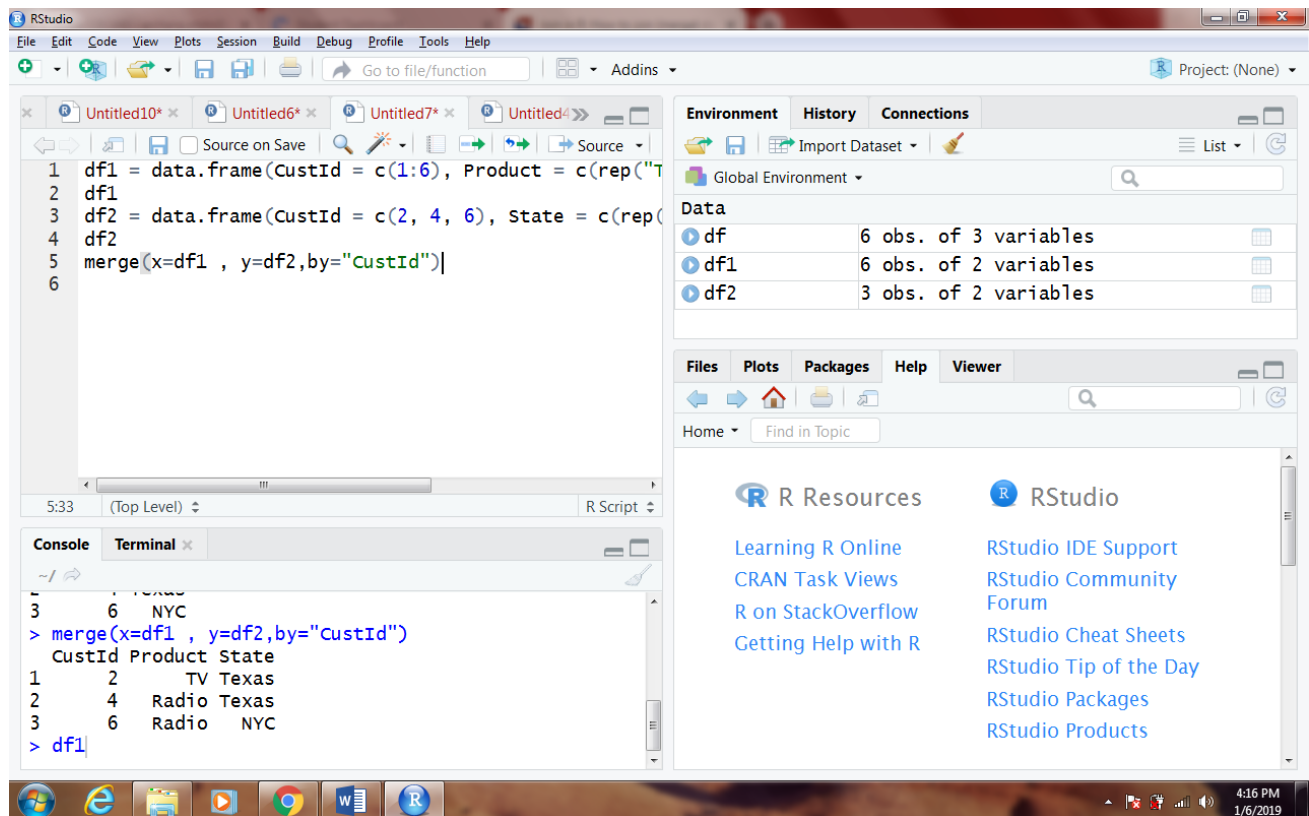
Not applicable.

5. 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
```

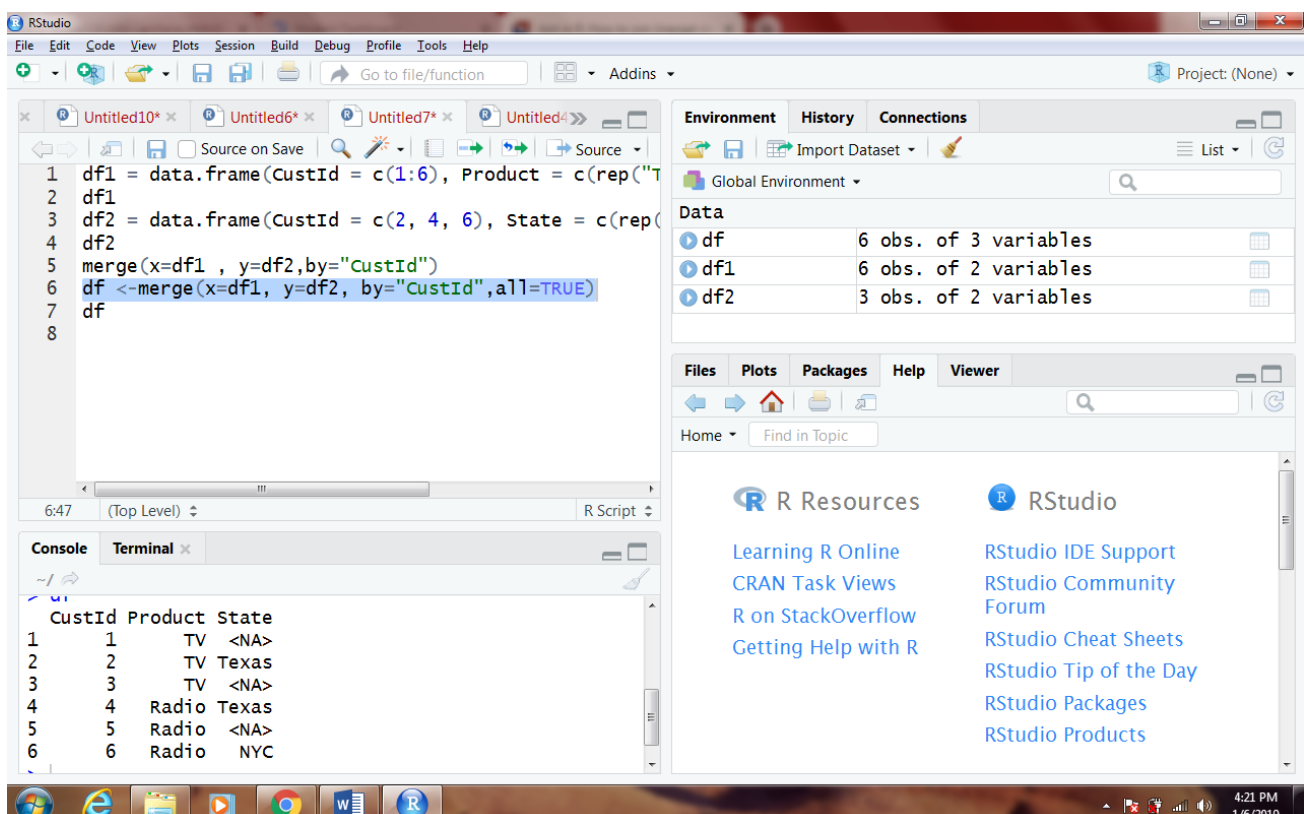For the above given data frames and tables perform the following operations:
- Return only the rows in which the left table have match.
- Returns all rows from both tables, join records from the left which have matching keys in the right table.
- Return all rows from the left table, and any rows with matching keys from the right table.
- Return all rows from the right table, and any rows with matching keys from the left table.

1.
```
df1 = data.frame(CustId = c(1:6), Product = c(rep("TV", 3), rep("Radio", 3)))
df1
df2 = data.frame(CustId = c(2, 4, 6), State = c(rep("Texas", 2), rep("NYC", 1)))
df2
merge(x=df1 , y=df2,by="CustId")
```

2.   df <-merge(x=df1, y=df2, by="CustId",all=TRUE)
df

3 df <-merge(x=df1 ,y=df2, by="CustId",all.x=TRUE)
df

3.   df <-merge(x=df1 ,y=df2 ,by="CustId",all.y=TRUE)
df

6. 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.

df <-merge(x=df1, y=df2, by=NULL)

df

```
library(dplyr)
semi_join(df1,df2)
```
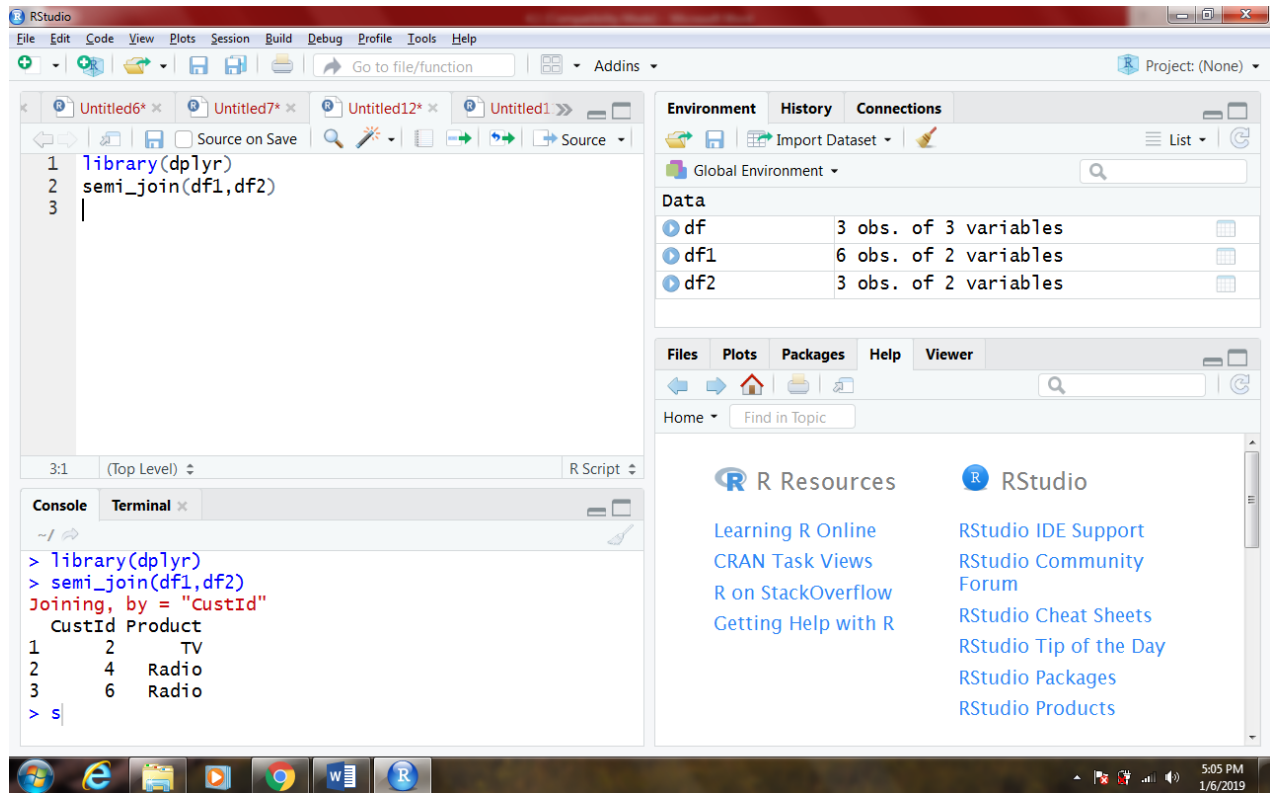


```
anti_join(df1,df2)
```

Untitled6* Untitled7* Untitled12* Untitled1

Source on Save

Source

```r
1  library(dplyr)
2  semi_join(df1,df2)
3  anti_join(df1,df2)
4
```

3:1   (Top Level)   R Script

Data

df	3 obs. of 3 variables	
df1	6 obs. of 2 variables	
df2	3 obs. of 2 variables	

Console Terminal

```
3     6    Radio
> anti_join(df1,df2)
Joining, by = "CustId"
  CustId Product
1      1      TV
2      3      TV
3      5    Radio
> s
```

5:06 PM
1/6/2019

Data Analytics

6. Expected Format

1. R file should be submitted where applicable.
2. R file should be in PDF or in .r format
3. Proper screenshots of the outputs should be submitted as well
4. The r codes, if submitted in any other format, will be subjected to deduction in marks

Note: Your solution will not be entertained if it is any other format, e.g., .zip, .doc, .rtf etc.

7. Approximate Time to Complete Task

20 mins.