**Week-2**

**Aim****: Write a R program to replace specific values in a column of Data Frame**

**Description:**

replace() function in[R Programming Language](https://www.geeksforgeeks.org/r-programming-language-introduction/) is used to replace the values in the specified string vector x with indices given in the list by those given in values.

**Syntax:** replace(list , position , replacement\_value)

We have four methods to replace a values in dataframe.

**1**.**Replacing NA values with 0’s**:

In this first we have to create a dataframe. The dataframe can be created using “c” function.and to identify the NA values we use “is.na ()”built in function if any NA is identified it can be replaced by 0.

**2. Replacing NA values with mean values:**

After creation of dataframe the dataframe can be assigned to another variable then it has a syntax based on syntax we can assign the values. In this we use mean function and in this na.rm=T,then it replace values.

**3.Replacing negative values with 0:**

After creation of dataframe negative values can be identified which is lessthan 0 in dataframe the values can be replaced by 0’s.

**4.Replacing a specific single value in dataframe:**

After creation of dataframe we use replace function and we give column name specific position which we have to change give the new value.

## Replace columns using mean() function

mean using a dataframe and [mean( ) function](https://www.geeksforgeeks.org/calculate-arithmetic-mean-in-r-programming-mean-function/). mean() function is used to calculate the arithmetic mean of the elements of the numeric vector passed to it as an argument.

**Syntax of mean() :**mean(x, trim = 0, na.rm = TRUE, …)

**Arguments:**

* x – any object
* trim – observations to be trimmed from each end of x before the mean is computed
* na.rm – TRUE to remove NA values

**Program:**

subject1 = c(80,85,NA,88,NA)

subject2=c(73,98, NA,74,60)

subject3=c(76,85,54,NA,80)

m=data.frame(subject1,subject2,subject3)

**Output:**

subject1 subject2 subject3

80 73 76

85 98 85

NA NA 54

88 74 NA

NA 60 80

**1.Replacing NA values with 0's:**

subject1 = c(80,85,NA,88,NA)

subject2=c(73,98, NA,74,60)

subject3=c(76,85,54,NA,80)

m=data.frame(subject1,subject2,subject3)

m[is.na(m)]=0

m

**Output:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| subject1 subject2 subject3  80 73 76  85 98 85  0 0 54  88 74 0  0 60 80  **2.Replacing NA values with mean values:**  subject1 = c(80,85,NA,88,NA)  subject2=c(73,98, NA,74,60)  subject3=c(76,85,54,NA,80)  m=data.frame(subject1,subject2,subject3)  r=m  r$subject1[is.na(r$subject1)]=mean(r$subject1,na.rm=T)  r  **Output:**  subject1 subject2 subject3  80.00000 73 76  85.00000 98 85  84.33333 NA 54  88.00000 74 NA  84.33333 60 80  **3.Replacing neagtive values with 0's:**  student1=c(-30,30,-40)  student2=c(-49,-50,60)  student3=c(80,-117,50)  q=data.frame(student1,student2,student3)  q[q<0]=0  q  **Output:**  student1 student2 student3  0 0 80  30 0 0  0 60 50  **4.Replace a single value and update in the table**  student1=c(30,30,40)  student2=c(49,50,60)  student3=c(80,27,50)  r=data.frame(student1,student2,student3)  r$subject1=replace(r$subject1,3,30)  r  **Output:**   |  | | --- | | student1 student2 student3  30 49 80  30 50 27  40 60 50 | |  | | |  | | --- | |  | | |
|  |
| |  | | --- | |  | |