How can I replace a character in a string using the -R command?

Authored by

stats writer

June 24, 2024

The -R command is a useful tool for replacing characters in a string. It allows you to specify a character or pattern to be replaced with another character or string. To use the -R command, you must first identify the string that you want to modify. Then, you can specify the character or pattern to be replaced and the replacement character or string. This command can be used in various scenarios, such as correcting typos or formatting errors in a text document, or updating data in a database. It is a simple and efficient way to make changes to a string without having to manually edit each occurrence.

How to replace a single character in a string on the R DataFrame column (find and replace)? To replace a first or all occurrences of a single character in a string use gsub(), sub(), str_replace(), str_replace_all() and functions from dplyr package of R. gsub() and sub() are R base functions and str_replace() and str_replace_all() are from the `stringr` package.

1. Quick Examples of Replace Character in a String

Following are quick examples of how to replace a character in a string column of R DataFrame.

```
# Quick Examples

# Example 1
# Replace first character occurrence in a string
df$address <- sub('f','F',df$address)

# Example 2
# Replace all characters occurrence in a string
df$work_address <- gsub('p','P',df$work_address)

# Example 3
# Replace first occurrence
library('stringr')
df$work_address <- str_replace(df$work_address,'P','p')

# Example 4
# Replace all occurrences
library('stringr')
df$address <- str_replace_all(df$address,'e','E')

# Example 5
# Replace first occurrence
library('dplyr')
df <- df %>%
```

mutate(address = str_replace(address, "E", "e"))

# Example 6
# Replace all occurrences
library('dplyr')
df <- df %>%
mutate(work_address = str_replace_all(work_address, "o", "O"))

let's create an R DataFrame and run these examples and explore the output.

```
# Create DataFrame
df <- data.frame(id=c(1,2,3,NA),
address=c('Orange St','Anton Blvd','Jefferson Pkwy',''),
work_address=c('Main St',NA,'Apple Blvd','Portola Pkwy'))
df
```

Output
id address work_address
#1 1 Orange St Main St
#2 2 Anton Blvd <NA>
#3 3 Jefferson Pkwy Apple Blvd
#4 NA Portola Pkwy

2. Using sub() - Replace Character in a String

`sub()` is a R Base function that is used to replace a specified character of first occurrences on a string (vector). This return a character vector of the same length and with the same attributes as the input column.

2.1 sub() Syntax

Following is the syntax of `sub()` function.

```
# Syntax of sub()
sub(pattern, replacement, x, ignore.case = FALSE, perl = FALSE,
fixed = FALSE, useBytes = FALSE)
```

2.2 Parameters

And the rest of the parameters are optional and they are set to default with a `False` value.

2.3 sub() Example - Replace Character in a String

sub() function is used to replace the first occurrence of a character with another character on a string column. Elements of input specified column which are not substituted will be returned unchanged.

```
# Replace first occurrence of a character
df$address <- sub('f','F',df$address)
print(df)
```

Output
id address work_address
#1 1 Orange St Main St
#2 2 Anton Blvd <NA>
#3 3 JeFferson Pkwy Apple Blvd
#4 NA Portola Pkwy

The result of the `sub()` function is assigned back to the same column (vector).

3. Use gsub() to Replace Character of all Occurrences in a String

`gsub()` is also R Base function used to replace all occurrences of the pattern character with another character in a string.

3.1 gsub() Syntax

Following is the syntax of `gsub()` function.

```
# Syntax of gsub()
gsub(pattern, replacement, x, ignore.case = FALSE, perl = FALSE,
fixed = FALSE, useBytes = FALSE)
```

3.2 Parameters

3.3 gsub() Example - Replace Character in a String

In the following example, replace all occurrences of character p (small letter p) with P (big letter P) on the `word_address` column of R DataFrame. The result of the `gsub()` function is assigned back to the same column (vector).

```
# Replace only first occurance of a character
df$work_address <- gsub('p','P',df$work_address)
print(df)
```

Output
id address work_address
1 1 Orange St Main St
2 2 Anton Blvd <NA>
3 3 JeFferson Pkwy APPle Blvd
4 NA Portola Pkwy

4. Use str_replace() to Replace Character in a String

str_replace() is a method from stringr package, `stringr` is a third-party package that provides a set of functions to work with strings as easily as possible. To use this, you need to load the library using `library("stringr")`. In case you don't have this package, install it using `install.packages("stringr")`.

It is used to replace a part of a string (character) on a column with another string or a character. You can also use pattern matching.

```
# Replace first occurrence
library('stringr')
df$work_address <- str_replace(df$work_address,'P','p')
df
```

Output
id address work_address
#1 1 Orange St Main St
#2 2 Anton Blvd <NA>
#3 3 JeFferson Pkwy ApPle Blvd
#4 NA portola Pkwy

5. Using str_replace_all() - Replace all Characters in a String

Use `str_replace_all()` method of `stringr` package to replace all occurrences of a character in a DataFrame column or a string.

In the following example, we update all occurrences of e with E on the address column.

```
# Replace all occurrences
library('stringr')
df$address <- str_replace_all(df$address,'e','E')
df
```

Output
id address work_address
#1 1 OrangE St Main St
#2 2 Anton Blvd <NA>
#3 3 JEFfErson Pkwy ApPle Blvd
#4 NA portola Pkwy

6. Using dplyr package

Let's use mutate() function from dplyr package to replace the first occurrence of a character in a string on R DataFrame. `dplyr` is a third-party package hence, you need to load the library using `library("dplyr")` to use its methods. In case you don't have this package, install it using `install.packages("dplyr")`.

For bigger data sets it is best to use the methods from `dplyr` package as they perform 30% faster. `dplyr` package uses C++ code to evaluate.

```
# Replace first occurrence
library('dplyr')
df <- df %>%
mutate(address = str_replace(address, "E", "e"))
print(df)
```

Output
id address work_address
#1 1 Orange St Main St
#2 2 Anton Blvd <NA>

#3 3 JeFfErson Pkwy ApPle Blvd
#4 NA portola Pkwy

Similarly use mutate() with str_replace_all() to replace all occurrences.

```
# Use mutate() with str_replace_all()
library('dplyr')
df <- df %>%
mutate(work_address = str_replace_all(work_address, "o", "O"))
print(df)
```

Output
id address work_address
#1 1 Orange St Main St
#2 2 Anton Blvd <NA>
#3 3 JeFfErson Pkwy ApPle Blvd
#4 NA pOrtOla Pkwy

Conclusion

In this article, you have learned how to replace the first and all occurrences of a character in a string. Learned gsub() and sub() are R base functions and str_replace() and str_replace_all() are from the stringr package which are used to find and replace.

Related Articles

References