# How can I use str_match in R, and what are some examples of its usage?

Authored by

**stats writer**

June 27, 2024

Str_match is a function in the programming language R that allows users to search for specific patterns within a string of characters. This function can be useful for data manipulation and analysis as it allows for the extraction of relevant information from a larger dataset. Examples of its usage include identifying and extracting specific words or phrases from a text, identifying and separating data based on a specific pattern, and cleaning and organizing data. Overall, str_match provides a powerful tool for data exploration and manipulation in R.

## Use str_match in R (With Examples)

## The str_match() function from the package in R can be used to extract matched groups from a string.

## This function uses the following syntax:

## str_match(string, pattern)

## where:

## string: Character vectorpattern: Pattern to look for

## The following examples show how to use this function in practice

### Example 1: Use str_match with Vector

## The following code shows how to use the str_match() function to extract matched patterns from a character vector:

```
library(stringr)

#create vector of strings
x <- c('Mavs', 'Cavs', 'Heat', 'Thunder', 'Blazers')

#extract strings that contain 'avs'
str_match(x, pattern='avs')

"avs"
"avs"
NA
NA
NA
```

The result is a matrix in which each row displays the matched pattern or an NA value if the pattern was not found.

For example:

The pattern 'avs' was found in the first element 'Mavs', so 'avs' was returned.The pattern 'avs' was found in the second element 'Cavs', so 'avs' was returned.The pattern 'avs was not found in the third element 'Heat' so NA was returned.

## And so on.

### Example 2: Use str_match with Data Frame

## Suppose we have the following data frame in R:

```
#create data frame
df <- data.frame(team=c('Mavs', 'Cavs', 'Heat', 'Thunder', 'Blazers'),
points=c(99, 104, 110, 103, 115))

#view data frame
df

team points
1 Mavs 99
2 Cavs 104
3 Heat 110
4 Thunder 103
5 Blazers 115
```



## The following code shows how to use the str_match() function to add a new column to the data frame that either does or does not contain a matched pattern for each team name:

```
library(stringr)

#create new column
df$match <- str_match(df$team, pattern='avs')

#view updated data frame
df

  team points match
1 Mavs     99   avs
2 Cavs    104   avs
3 Heat    110  <NA>
4 Thunder 103  <NA>
5 Blazers 115  <NA>
```

The new column titled match contains either the pattern 'avs' or NA, depending on whether the pattern is found in the team column.