

How can I extract a substring starting from the end of a string using the R programming language?

Authored by
stats writer

June 24, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I extract a substring starting from the end of a string using the R programming language?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=151390>

In the R programming language, the function "substr()" can be used to extract a substring from a given string starting from a specified position. To extract a substring from the end of a string, the "nchar()" function can be used to determine the length of the string, and then the "substr()" function can be applied with the starting position set as the length of the string minus the desired number of characters. This will result in a substring starting from the end of the original string.

R: Extract Substring Starting from End of String

You can use the following methods to extract a substring in R starting from the end of the string:

Method 1: Use Base R

```
#define function to extract n characters starting from end
```

```
substr_end <- function(x, n){  
  substr(x, nchar(x)-n+1, nchar(x))  
}
```

```
#extract 3 characters starting from end
```

```
substr_end(my_string, 3)
```

Method 2: Use stringr Package

```
library(stringr)
```

```
#extract 3 characters starting from end
```

```
str_sub(my_string, start = -3)
```

Both of these examples extract the last three characters from the string called my_string.

The following examples show how to use each method in practice with the following data frame:

```
#create data frame
```

```
df <- data.frame(team=c('Mavericks', 'Lakers', 'Hawks',  
'Nets', 'Warriors'),  
points=c(100, 143, 129, 113, 123))
```

```
#view data frame
```

```
df
```

```
team points
```

```
1 Mavericks 100
```

```
2 Lakers 143
```

```
3 Hawks 129
```

```
4 Nets 113
```

```
5 Warriors 123
```

Example 1: Extract Substring Starting from End Using Base R

The following code shows how to define a custom function in base R and then use the function to extract the last three characters from each string in the team

column:

#define function to extract n characters starting from end

```
substr_end <- function(x, n){  
  substr(x, nchar(x)-n+1, nchar(x))  
}
```

#create new column that extracts last 3 characters from team column

```
df$team_last3 <- substr_end(my_string, 3)
```

#view updated data frame

df

team points team_last3

1 Mavericks 100 cks

2 Lakers 143 ers

3 Hawks 129 wks

4 Nets 113 ets

5 Warriors 123 ors

Notice that the new column called team_last3 contains the last three characters of each string in the team column of the data frame.

Example 2: Extract Substring Starting from End Using stringr Package

The following code shows how to use the `str_sub()` function from the `stringr` package in R to extract the last three characters from each string in the `team` column:

```
library(stringr)
#create new column that extracts last 3 characters from
team column
df$team_last3 <- str_sub(df$team, start = -3)

#view updated data frame
df
```

	team	points	team_last3
1	Mavericks	100	cks
2	Lakers	143	ers
3	Hawks	129	wks
4	Nets	113	ets
5	Warriors	123	ors

Notice that the new column called `team_last3` contains the last three characters of each string in the `team` column of the data frame.

This matches the results from the previous method using base R.

ARABPSYCHOLOGY.COM