

How can I use dplyr to remove the first character from strings?

Authored by
stats writer

June 24, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use dplyr to remove the first character from strings?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=151406>

Dplyr is a data manipulation package in R that allows for efficient and streamlined data manipulation. One of its useful functions is the ability to remove the first character from strings. This can be achieved by using the "str_sub" function, which allows for the selection of specific characters from a string. By specifying a start position of 2, the first character can be excluded, effectively removing it from the string. This can be applied to a single string or to a column of strings in a data frame, making it a versatile tool for data cleaning and manipulation.

Remove First Character from Strings Using dplyr

You can use the following basic syntax in to remove the first character from each string in a particular column:

```
library(dplyr)
```

```
df_new <- df %>% mutate(across(c('my_column'),  
substr, 2, nchar(my_column)))
```

This particular syntax removes the first character from each string in the column called my_column.

Note that we use the substr() function to extract the substring ranging from the second character in each string to the length of the string.

This has the effect of removing the first character from the string.

The following example shows how to use this syntax in

practice.

Example: Remove First Character from Strings Using dplyr

Suppose we have the following data frame in R:

```
#create data frame
```

```
df <- data.frame(team=c('XMavs', 'XPacers', 'XHawks',  
'XKings', 'XNets', 'XCeltics'),  
points=c(104, 110, 134, 125, 114, 124))
```

```
#view data frame
```

```
df
```

```
team points
```

```
1 XMavs 104
```

```
2 XPacers 110
```

```
3 XHawks 134
```

```
4 XKings 125
```

```
5 XNets 114
```

```
6 XCeltics 124
```

Suppose we would like to remove the first character from each string in the team column.

We can use the following syntax to do so:

```
library(dplyr)
```

```
#remove first character from each string in 'team'  
column
```

```
df_new <- df %>% mutate(across(c('team'), substr, 2,  
nchar(team)))
```

```
#view updated data frame
```

```
df_new
```

```
team points
```

```
1 Mavs 104
```

```
2 Pacers 110
```

```
3 Hawks 134
```

```
4 Kings 125
```

```
5 Nets 114
```

```
6 Celtics 124
```

Notice that the first character of each string in the team column has been removed.

Note that the `nchar()` function is used to calculate the total number of characters in a string.

Thus, we use the `substr()` function to extract the

substring ranging from the second character to the last character in each string, which is equivalent to removing the first character from each string.

Note: If you'd like to remove the first character from strings in multiple columns, simply include multiple column names in the across() function.

ARABPSYCHOLOGY.COM