

How can I use the “NOT IN” operator in R? Can you provide some examples?

Authored by
stats writer

May 4, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use the “NOT IN” operator in R? Can you provide some examples?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=142578>

The "NOT IN" operator in R is a logical operator that allows users to specify a set of values to exclude from a data frame, vector, or other data structure. It is particularly useful for filtering out specific elements or categories from a dataset. To use the "NOT IN" operator, the user must first define the set of values to be excluded and then use the operator in combination with the ">" or "<" symbol. For example, "variable NOT IN c(1,2,3)" will exclude the values 1, 2, and 3 from the variable. This can be applied in various scenarios, such as removing outliers or irrelevant data points from an analysis.

Use "NOT IN" Operator in R (With Examples)

You can use the following basic syntax to select all elements that are not in a list of values in R:

```
!(data %in% c(value1, value2, value3, ...))
```

The following examples show how to use this syntax in practice.

Example 1: How to Use "NOT IN" with Vectors

The following code shows how to select all values in a vector in R that are not in a certain list of values:

```
#define numeric vector
```

```
num_data <- c(1, 2, 3, 3, 4, 4, 5, 5, 6)
```

```
#display all values in vector not equal to 3 or 4
```

```
num_data
```

1 2 5 5 6

All values that are not equal to 3 or 4 are shown in the output.

Note that we can use the same syntax to select all elements in a vector that are not in a certain list of characters:

```
#define vector of character data
char_data <- c('A', 'A', 'A', 'B', 'B', 'C', 'C', 'D', 'D', 'D')

#display all elements in vector not equal to 'A', or 'C'
char_data

"B" "B" "D" "D" "D"
```

All values that are not equal to 'A' or 'C' are shown in the output.

Example 2: How to Use "NOT IN" with Data Frames

The following code shows how to select all rows in a data frame in R in which a certain column is not equal to certain values:

```
#create data frame
```

```
df <- data.frame(team=c('A', 'A', 'B', 'B', 'C', 'C', 'D'),  
points=c(77, 81, 89, 83, 99, 92, 97),  
assists=c(19, 22, 29, 15, 32, 39, 14))
```

```
#view data frame
```

```
df
```

```
team points assists
```

```
1 A 77 19
```

```
2 A 81 22
```

```
3 B 89 29
```

```
4 B 83 15
```

```
5 C 99 32
```

```
6 C 92 39
```

```
7 D 97 14
```

```
#select all rows where team is not equal to 'A' or 'B'
```

```
subset(df, !(team %in% c('A', 'B')))
```

```
team points assists
```

```
5 C 99 32
```

```
6 C 92 39
```

```
7 D 97 14
```

Notice that all rows that do not have an 'A' or 'B' in the team column are returned.

We can also use similar syntax to select all rows in which a certain column is not equal to certain numeric values:

```
#create data frame
```

```
df <- data.frame(team=c('A', 'A', 'B', 'B', 'C', 'C', 'D'),  
points=c(77, 81, 89, 83, 99, 92, 97),  
assists=c(19, 22, 29, 15, 32, 39, 14))
```

```
#view data frame
```

```
df
```

```
team points assists
```

```
1 A 77 19
```

```
2 A 81 22
```

```
3 B 89 29
```

```
4 B 83 15
```

```
5 C 99 32
```

```
6 C 92 39
```

```
7 D 97 14
```

```
#select all rows where team is not equal to 'A' or 'B'
```

```
subset(df, !(points %in% c(89, 99)))
```

```
team points assists
```

```
1 A 77 19
```

```
2 A 81 22
```

```
4 B 83 15
```

```
6 C 92 39
```

```
7 D 97 14
```

Notice that all rows that are not equal to 89 or 99 in the points column are returned.

ARABPSYCHOLOGY.COM