

How can we use a “NOT IN” operator in SAS to perform a query?

Authored by
stats writer

June 25, 2024

RECOMMENDED CITATION

stats writer (2024). *How can we use a “NOT IN” operator in SAS to perform a query?*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=152891>

The "NOT IN" operator in SAS is used to perform a query that excludes specific values from a dataset or table. It is a logical operator that allows us to specify a list of values that should not be included in the query results. This operator is particularly useful when we want to filter out data that does not meet certain criteria or when we want to compare two datasets and identify records that are not present in both. By using the "NOT IN" operator, we can easily and efficiently retrieve the desired information from our datasets in SAS.

SAS: Use a "NOT IN" Operator

You can use the NOT IN operator in SAS to return only the rows where a variable does not have a value in some list of values.

The following example shows how to use the NOT IN operator in practice.

Example: Using NOT IN Operator in SAS

Suppose we have the following dataset in SAS that contains information about various basketball players:

```
/*create dataset*/  
data my_data;  
input team $ points;  
datalines;  
Cavs 12  
Cavs 14  
Warriors 15
```

Hawks 18

Mavs 31

Mavs 32

Mavs 35

Celtics 36

Celtics 40

;

run;

/*view dataset*/

proc printdata=my_data;

Obs	team	points
1	Cavs	12
2	Cavs	14
3	Warriors	15
4	Hawks	18
5	Mavs	31
6	Mavs	32
7	Mavs	35
8	Celtics	36
9	Celtics	40

We can use the NOT IN operator in PROC SQL to select only the rows where the team is not equal to 'Cavs' or 'Celtics':

```
/*select all rows where team is not 'Cavs' or 'Celtics'*/  
proc sql;  
select *  
from my_data  
where team not in ('Cavs', 'Celtics');quit;
```

team	points
Warriors	15
Hawks	18
Mavs	31
Mavs	32
Mavs	35

Notice that the only rows returned are the ones where the team is not equal to 'Cavs' or 'Celtics.'

You can also use the NOT IN operator within a SET statement to create a new dataset that only contains rows where the team is not equal to 'Cavs' or 'Celtics':

```
/*create new dataset that only contains rows where  
team is not Cavs or Celtics*/  
data new_data;  
set my_data;  
where team not in ('Cavs', 'Celtics');
```

```
run;
```

```
/*view new dataset*/
```

```
proc printdata=new_data;
```

Obs	team	points
1	Warriors	15
2	Hawks	18
3	Mavs	31
4	Mavs	32
5	Mavs	35

The new dataset called new_data only contains the rows from the original dataset where the team is not equal to 'Cavs' or 'Celtics.'

The following tutorials explain how to perform other common tasks in SAS: