How can the SAS SAS procedure SQL be utilized to implement the SAS SAS procedure SQL by incorporating the UNION operator?

Authored by

stats writer

June 25, 2024

The SAS procedure SQL, a powerful tool for data manipulation and analysis, can be further enhanced by incorporating the UNION operator. This operator allows for the combination of data from multiple sources, resulting in a single dataset that contains all relevant information. By utilizing the SAS procedure SQL and incorporating the UNION operator, users can efficiently merge and analyze data from different tables, databases, or even different software programs. This functionality provides a streamlined approach to data integration, allowing for more efficient and accurate analysis in the SAS environment.

SAS: Use UNION in PROC SQL

You can use the UNION operator in the PROC SQL statement in SAS to combine two datasets vertically.

The following example shows how to use the UNION operator in practice.

Example: Using UNION in PROC SQL in SAS

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

```
/*create first dataset*/
data data1;
input team $ points;
datalines;
A 12
A 14
A 15
```

```
A 18
A 20
A 22
;
run;

/*view first dataset*/
proc printdata=data1;
```

Obs	team	points
1	A	12
2	A	14
3	A	15
4	A	18
5	A	20
6	A	22

And suppose we have another dataset in SAS that also contains information about various basketball players:

```
/*create second dataset*/
data data2;
input team $ points;
datalines;
A 12
A 14
```

B 23

B 25

B 29

B 30

;

run;

/*view second dataset*/

proc printdata=data2;

Obs	team	points
1	A	12
2	A	14
3	B	23
4	B	25
5	B	29
6	B	30

We can use the UNION operator in the PROC SQL statement to combine these two datasets vertically and only keep the unique rows:

/*combine tables vertically and only keep unique rows*/

proc sql;

title 'data1 UNION data2';

select * from data1

unionselect * from data2;

quit;

data1 UNION data2

team	points
A	12
A	14
A	15
A	18
A	20
A	22
B	23
B	25
B	29
B	30

Notice that the two datasets have been combined vertically and only the unique rows are kept.

We can also use the UNION ALL operator in the PROC SQL statement to combine these two datasets vertically and keep *all* of the rows:

/*combine tables vertically and keep all rows*/

proc sql;

title 'data1 UNION ALL data2';

select * from data1

```
union allselect * from data2;
quit;
```

data1 UNION ALL data2

team	points
A	12
A	14
A	15
A	18
A	20
A	22
A	12
A	14
B	23
B	25
B	29
B	30

Notice that the two datasets have been combined vertically and all rows are kept from both datasets, even the ones that are duplicates.

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