How can I count distinct values in SAS and provide examples?

Authored by

stats writer

July 1, 2024

SAS is a statistical software that allows users to analyze and manipulate large datasets. One common task in data analysis is to count the number of distinct values in a variable or column. This can be easily achieved in SAS using the PROC FREQ procedure.

To count distinct values in SAS, the user can specify the variable or column of interest in the PROC FREQ statement, along with the keyword "nlevels" which stands for number of levels. This will generate a table with the distinct values and their corresponding frequencies.

For example, if we have a dataset containing customer information and we want to count the number of distinct cities in which our customers reside, we can use the following SAS code:

PROC FREQ data = customer_data;
TABLES city / nlevels;
RUN;

This will produce a table with the distinct cities and their corresponding frequencies. For instance, if our customers reside in New York, Chicago, and Los Angeles, the table will display these three cities with their respective frequencies.

In conclusion, counting distinct values in SAS can be easily achieved using the PROC FREQ procedure. This functionality is useful for gaining insights into the distribution of data and identifying any potential outliers.

Count Distinct Values in SAS (With Examples)

You can use the following methods to count distinct values in a dataset in SAS:

Method 1: Count Distinct Values in One Column

```
proc sql;
select count(distinct var1) as distinct_var1
from my_data;
quit;
```

Method 2: Count Distinct Values by Group

```
proc sql;
select var1, count(distinct var2) as distinct_var2
from my_data
group by var1;
```

quit;

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

```
/*create dataset*/
data my_data;
input team $ points;
datalines;
Mavs 10
Mavs 13
Mavs 13
Mavs 15
Mavs 15
Rockets 9
Rockets 10
Rockets 10
Spurs 18
Spurs 19
;
run;

/*view dataset*/
proc printdata=my_data;
```

Obs	team	points
1	Mavs	10
2	Mavs	13
3	Mavs	13
4	Mavs	15
5	Mavs	15
6	Rockets	9
7	Rockets	10
8	Rockets	10
9	Spurs	18
10	Spurs	19

Example 1: Count Distinct Values in One Column

The following code shows how to count the total distinct values in the team column:

```
/*count distinct values in team column*/
proc sql;
select count(distinct team) as distinct_teams
from my_data;
quit;
```

distinct_teams
3

From the output we can see that there are **3** distinct values in the team column.

We can confirm this manually by observing that there are three different teams: Mavs, Rockets, and Spurs.

Example 2: Count Distinct Values by Group

The following code shows how to count the distinct values in the **points** column, grouped by the **team** column:

```
/*count distinct values in points column, grouped by team*/
proc sql;
select team, count(distinct points) as distinct_points
from my_data
group by team;
quit;
```

team	distinct_points
Mavs	3
Rockets	2
Spurs	2

The resulting table shows the number of distinct values in the points column, grouped by each of

the teams.

Additional Resources

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