How do you create frequency tables in SAS? Can you provide examples?

Authored by

stats writer

July 1, 2024

Frequency tables in SAS are used to display the count or percentage of occurrences for each unique value in a dataset. To create a frequency table in SAS, the PROC FREQ procedure is used. This procedure allows you to specify the variables you want to include in the table and the desired statistics such as count or percentage. Here is an example of creating a frequency table for the variable "gender" in a dataset called "survey":

```
PROC FREQ data=survey;
TABLES gender;
RUN;
```

This will generate a table with the count and percentage of males and females in the survey dataset. Additionally, you can use the "out" option to save the frequency table as a new dataset for further analysis. Frequency tables are a useful tool for summarizing categorical data and can provide valuable insights in data analysis.

Create Frequency Tables in SAS (With Examples)

You can use proc freq in SAS to quickly create frequency tables for one or more variables in a dataset.

The following examples show how to use this procedure with the SAS built-in dataset called , which contains various characteristics for 100,000 mothers that recently gave birth.

We can use proc print to view the first 10 observations from this dataset:

```
/*view first 10 observations from BirthWgt dataset*/
proc printdata=sashelp.BirthWgt (obs=10);
```

```
run;
```

Obs	LowBirthWgt	Married	AgeGroup	Race	Drinking	Death	Smoking	SomeCollege
1	No	No	3	Asian	No	No	No	Yes
2	No	No	2	White	No	No	No	No
3	Yes	Yes	2	Native	No	Yes	No	No
4	No	No	2	White	No	No	No	No
5	No	No	2	White	No	No	No	Yes
6	No	No	2	White	No	No	No	
7	No	No	2	Asian	No	No	No	Yes
8	No	No	3	White	No	No	No	Yes
9	No	Yes	1	Black	No	No	No	No
10	No	No	2	Native	No	No	No	Yes

Example 1: Frequency Table for One Variable

We can use the following code to create a frequency table for the Race variable:

```
/*create frequency table for Race variable*/
proc freqdata=sashelp.BirthWgt;
tables Race;
run;
```

The FREQ Procedure

Race	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Asian	5224	5.22	5224	5.22
Black	14133	14.13	19357	19.36
Hispanic	22139	22.14	41496	41.50
Native	942	0.94	42438	42.44
White	57562	57.56	100000	100.00

The output table contains four columns:

Frequency: The total number of observations that fell in a certain category.Percent: The percentage of total observations that fell in a certain category.Cumulative Frequency: The total number of observations that have been accounted for up to and including the current row.Cumulative Percent: The cumulative percentage of total observations that have been accounted for up to and including the current row.

For example, from the output table we can see:

The total number of Hispanic mothers was 22,139.The percentage of total mothers who were Hispanic was 22.14%.The total number of mothers who were Asian, Black, or Hispanic was 41,496.The cumulative

percentage of mothers who were Asian, Black, or Hispanic was 41.50%.

Example 2: Frequency Table for One Variable (Sorted)

By default, frequency tables are sorted in alphabetical order based on the category names. However, we can use the order function to sort the categories by frequency instead:

```
/*create frequency table for Race variable, sorted by frequency*/
proc freqdata=sashelp.BirthWgt order=freq;
tables Race;
run;
```

The FREQ Procedure

Race	Frequency	Percent	Cumulative Frequency	Cumulative Percent
White	57562	57.56	57562	57.56
Hispanic	22139	22.14	79701	79.70
Black	14133	14.13	93834	93.83
Asian	5224	5.22	99058	99.06
Native	942	0.94	100000	100.00

Example 3: Frequency Table for One Variable (Include Missing Values)

By default, missing values are not included in

frequency tables.

However, we can use the missing command to tell SAS to include a row to count the frequency of missing values:

```
/*create frequency table for Race variable, sorted by frequency*/
proc freqdata=sashelp.BirthWgt order=freq;
tables Race / missing;
run;
```

The FREQ Procedure

Race	Frequency	Percent	Cumulative Frequency	Cumulative Percent
White	57562	57.56	57562	57.56
Hispanic	22139	22.14	79701	79.70
Black	14133	14.13	93834	93.83
Asian	5224	5.22	99058	99.06
Native	942	0.94	100000	100.00

Since no additional row was added to the frequency table, this tells us that there were no missing values for Race in the original dataset.

Example 4: Frequency Table for Multiple Variables

To create a frequency table for multiple variables at once, we can simply include multiple variable names in the tables argument.

For example, we can use the following code to create a frequency table for both Race and AgeGroup:

```
/*create frequency table for Race and AgeGroup variables, both sorted by frequency*/
proc freqdata=sashelp.BirthWgt order=freq;
tables Race AgeGroup;
run;
```

The FREQ Procedure

Race	Frequency	Percent	Cumulative Frequency	Cumulative Percent
White	57562	57.56	57562	57.56
Hispanic	22139	22.14	79701	79.70
Black	14133	14.13	93834	93.83
Asian	5224	5.22	99058	99.06
Native	942	0.94	100000	100.00

AgeGroup	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2	75633	75.63	75633	75.63
3	14122	14.12	89755	89.76
1	10245	10.25	100000	100.00

We can see that a frequency table was created for both variables.