How can cases be weighted?

Authored by

stats writer

June 23, 2024

Cases can be weighted by assigning numerical values to different factors that contribute to the outcome of a case. These factors can include the complexity of the case, the potential impact of the outcome, and the resources required to resolve the case. By assigning weights to these factors, a more accurate representation of the importance and complexity of each case can be determined, allowing for fairer and more effective decision-making. This weighting process can be applied to various fields such as legal cases, research studies, and statistical analysis.

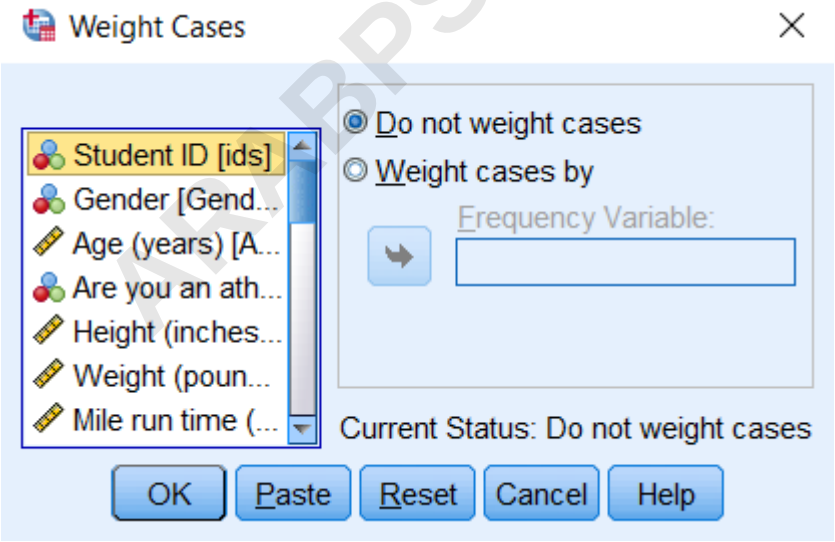
Weighting Cases

In SPSS, *weighting cases* allows you to assign "importance" or "weight" to the cases in your dataset. Some situations where this can be useful include:

Your data is in the form of counts (the number of occurrences) of factors or events. The "weight" is the number of occurrences.Your data requires adjustments to correct for over- or under-representation of certain characteristics in your sample. (This often happens with large surveys: a "weighting" variable is developed to adjust a sample's composition to be reflective of the population's composition, or to control for over- or under-reporting from a certain group.)
The Pew Research Center often makes their raw survey data available online to the public. Many of these surveys include weighting as a part of the study methodology.

Weighting cases in SPSS works the same way for both situations.

To turn on case weights, click **Data > Weight Cases**.



To enable a weighting variable, click **Weight cases by**, then double-click on the name of the weighting variable in the left-hand column to move it to the **Frequency Variable** field. Click **OK.**To turn off an enabled weighting variable, open Weight Cases window again, and click **Do not weight**

cases. Click **OK.**

Example: Reproducing an Existing 3x2 Table

Problem Statement

Suppose you are helping a friend with their statistics homework, and see that they have included the following write-up in their report:

John Smith
Statistics 101

Class Rank * Have you picked a major yet? Crosstabulation

Count

		Have you picked a major yet?		
		No	Yes	Total
Class Rank	Freshman	212	114	326
	Junior	92	198	290
	Sophomore	171	168	339
Total		475	480	955

The p-value is less than .05 which proves that most people haven't picked a major yet

You immediately notice several things wrong with this report so far:

You can never *prove* anything with a hypothesis test--not even if the p-value is really, really small.You do not see any test statistics anywhere, and it is unclear what test he has run.

You get the feeling that they may have used the Chi-square test of independence, but want to verify this for yourself.

Running the Procedure

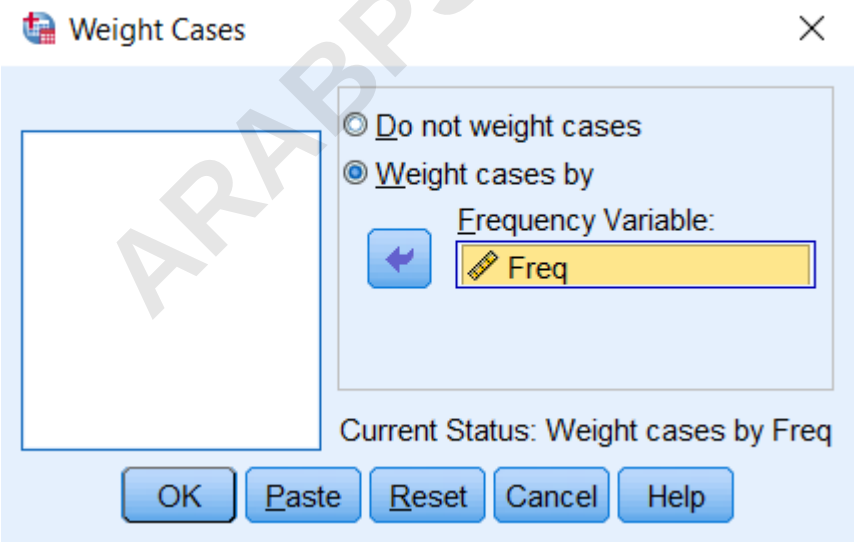
Whenever you want to re-create a frequency table or crosstab, you first need to figure out how many unique combinations of the factors there are, and how many observations there were for each factor combination. In this situation, we have two variables: *ClassRank* (which has three levels) and *PickedAMajor* (which has two levels). So there are 3*2 = 6 unique factor combinations. They are:

Freshman	No	212 responses
Freshman	Yes	114 responses
Sophomore	No	171 responses
Sophomore	Yes	168 responses
Junior	No	92 responses
Junior	Yes	198 responses

When we go to enter our data in SPSS, we will need to create three new variables: ClassRank, PickedAMajor, and a frequency variable (let's name it "Freq"). After entering the data, your Data View window should look like this:

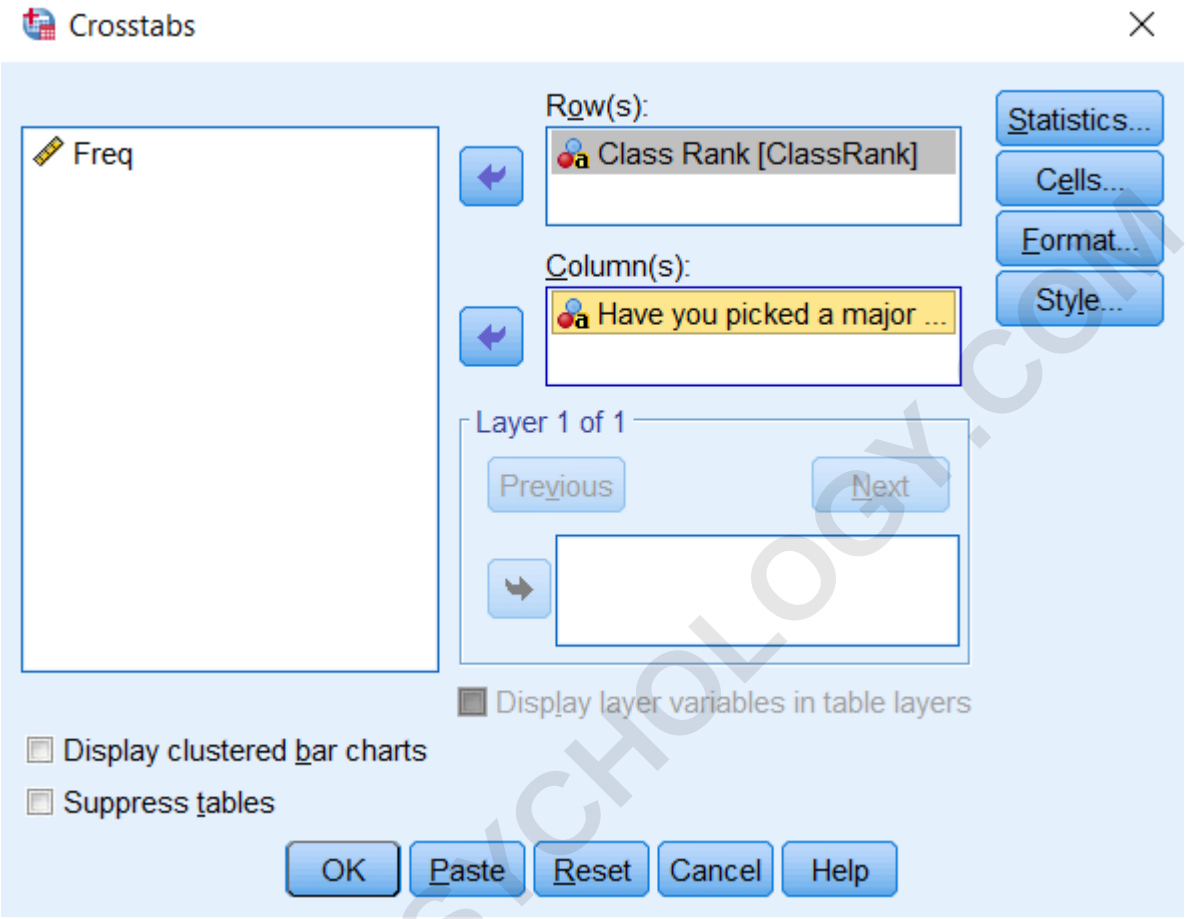
	ClassRank	PickedAMajor	Freq
1	Freshman	No	212
2	Freshman	Yes	114
3	Sophomore	No	171
4	Sophomore	Yes	168
5	Junior	No	92
6	Junior	Yes	198

Now we need to weight the cases with respect to Freq. Click **Data > Weight Cases**.



Click **Weight cases by**, then double-click Freq to move it to the **Frequency Variable** field. Click **OK**.

Now we can run our crosstab and verify your friend's results. Click **Analyze > Descriptive Statistics > Crosstabs**.



When the Crosstabs window opens, select the variable ClassRank in the left column and move it to the **Row(s)** field using the first arrow button, then select variable PickedAMajor in the left column and transfer it to the **Column(s)** field using the second arrow button. Doing this will reproduce the 3x2 table that your friend made.

To produce a Chi-square test of independence, click **Statistics**. This will open the *Crosstabs:* Statistics window. Select the **Chi-square** check box in the upper left-hand corner, then click **Continue**.

Click **OK** to run the crosstab.

Syntax

```
WEIGHT BY Freq.
```

CROSSTABS

```
/TABLES=ClassRank BY PickedAMajor
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ
/CELLS=COUNT
/COUNT ROUND CELL.
```

Output

Within your output, you should see the following two tables:

Class Rank * Have you picked a major yet? Crosstabulation

Count

		Have you picked a major yet?		
		No	Yes	Total
Class Rank	Freshman	212	114	326
	Junior	92	198	290
	Sophomore	171	168	339
Total		475	480	955

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	68.207[a]	2	.000
Likelihood Ratio	69.578	2	.000
N of Valid Cases	955		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 144.24.

The 3x2 table matches your friend's output. From the Chi-square Tests table, we see that this test result was significant at the 5% level ($\chi(2)=68.207$, $p < .001$). From this result, we infer that there is a significant association between a student's class rank and whether or not they have picked a major.

For more information

CrosstabsChi-square Test of Independence