Frequency in StatCrunch Transcript

TutorTube: UNT Learning Center

1 Introduction

Hello! In this video, we will be finding relative, cumulative, and relative cumulative frequency in StatCrunch for both raw data and summary data. Finding these frequencies by hand can be very time consuming, but luckily, StatCrunch lets us calculate them much more quickly.

1.1 Summary Data

So first, let's start with summary data. If data is presented in summary form, this means that it is already separated into bins and organized into a frequency table, as we can see here. Frequency tables are a method of summarizing categorical (or qualitative) data. And here we have qualitative data in the form of the category names, and then the frequency (or count) of each item. So for example, we know from this table that there were 589 items in the D category, and 917 items in the E category. The question then asks us to find relative frequency, cumulative frequency, and relative cumulative frequency for the data.

Category	Frequency
A	1131
B	1087
C	1362
D	589
E	917

1.1.1 Relative Frequency

In order to do this by hand, we would have to find the total count of items, and then apply the formulas for relative frequency and cumulative frequency multiple times. StatCrunch will do this for us very quickly.
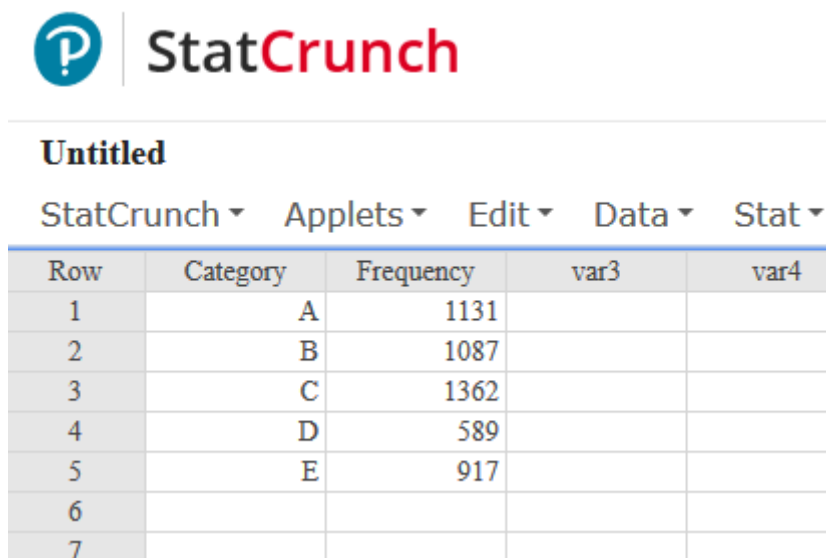

Figure 1: Summary Data in StatCrunch

Once we have your data entered in StatCrunch like this, we click on "Data," then "Compute," then "Multiple Expressions."
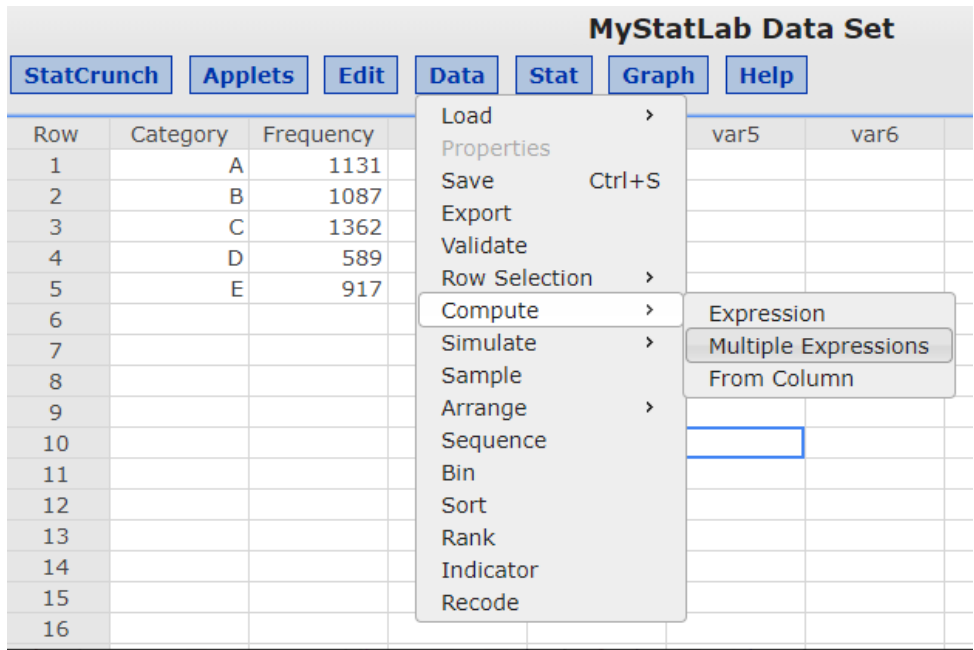
Figure 2: Select "Data", "Compute", "Multiple Expressions"

First, let's calculate the relative frequency. Naming your expressions here is optional, but I'll call the first one "relative freq." Now we just click on "Build."
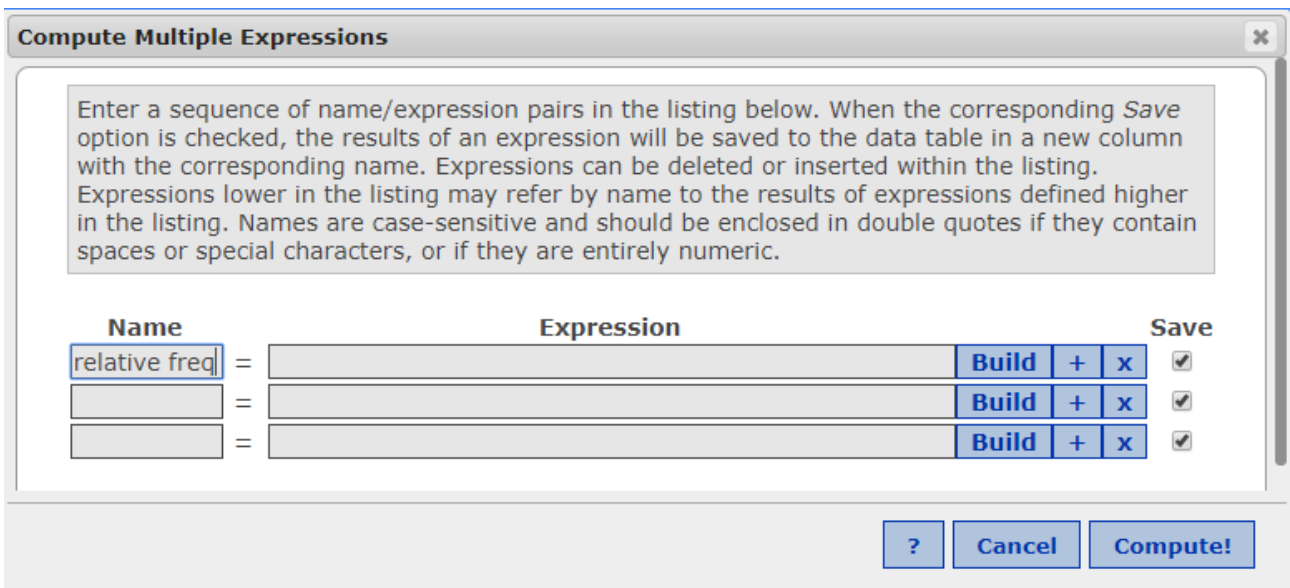

Figure 3: The Compute Multiple Expressions Window

It will open up this window. Now we enter the formula for relative frequency. Recall that formula for relative frequency is:

$$\text{Relative Frequency} = \frac{\text{Frequency}}{\text{Sum of All Frequencies}}$$

And in StatCrunch, this is created by double clicking on the name of the column that contains your frequency numbers (here this is called Frequency). Then click or type the division symbol. And now we scroll down in the functions box until you find the "sum" function. Double click on that. Now find your column name again, and double click again. Then hit "Okay."
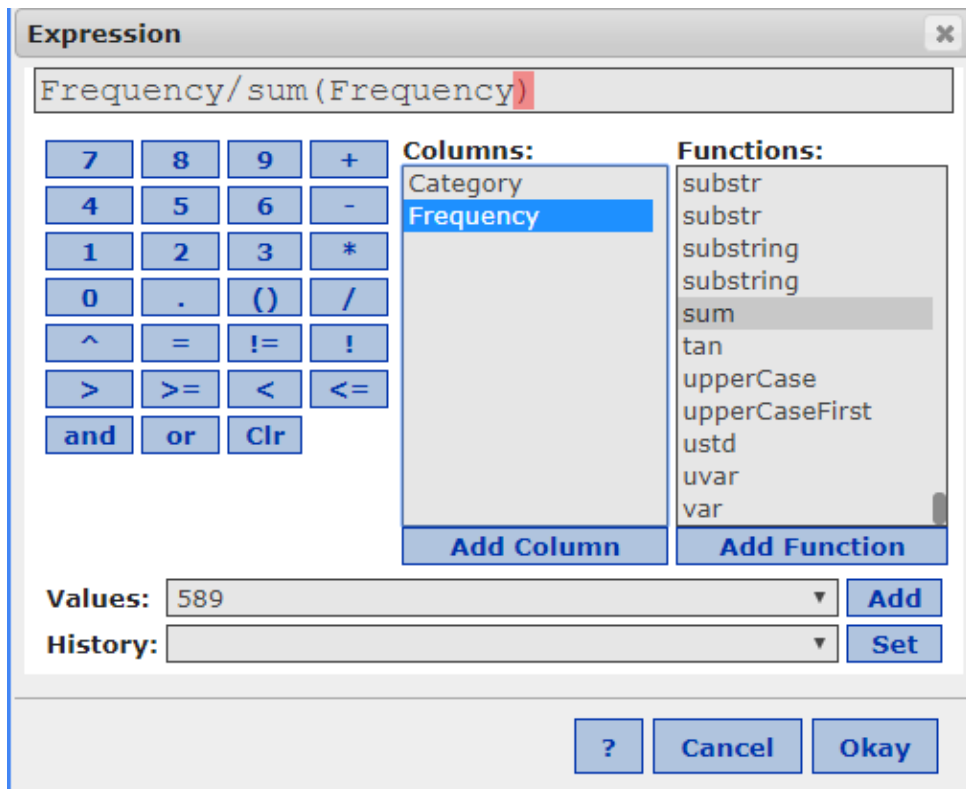
Figure 4: The Expression Window: Enter "Frequency/sum(Frequency)"

Ok, so we are taken back to this screen. And you can see it entered our expression into the window. Now we just hit "Compute" and notice that we now have a new column called relative frequency with all of our results automatically calculated. You just have to be sure to round your numbers to the correct number of decimals when giving your answers.

Figure 5: Relative Frequency Results

1.1.2 Cumulative Frequency

Next let's find Cumulative Frequency. Again, click on "Data," then "Compute," then "Multiple Expressions." This time I'll call the expression "cumu freq" and again we click on "Build."

And recall that cumulative frequency is found by adding each of the frequency values in each row. And StatCrunch has a function that does this automatically. So look in the functions box again, but this time we are going to look for a function called "cumsum." Double click on that. Now double click on "Frequency" again, this tells StatCrunch to find the cumulative sum of all of our frequency values.
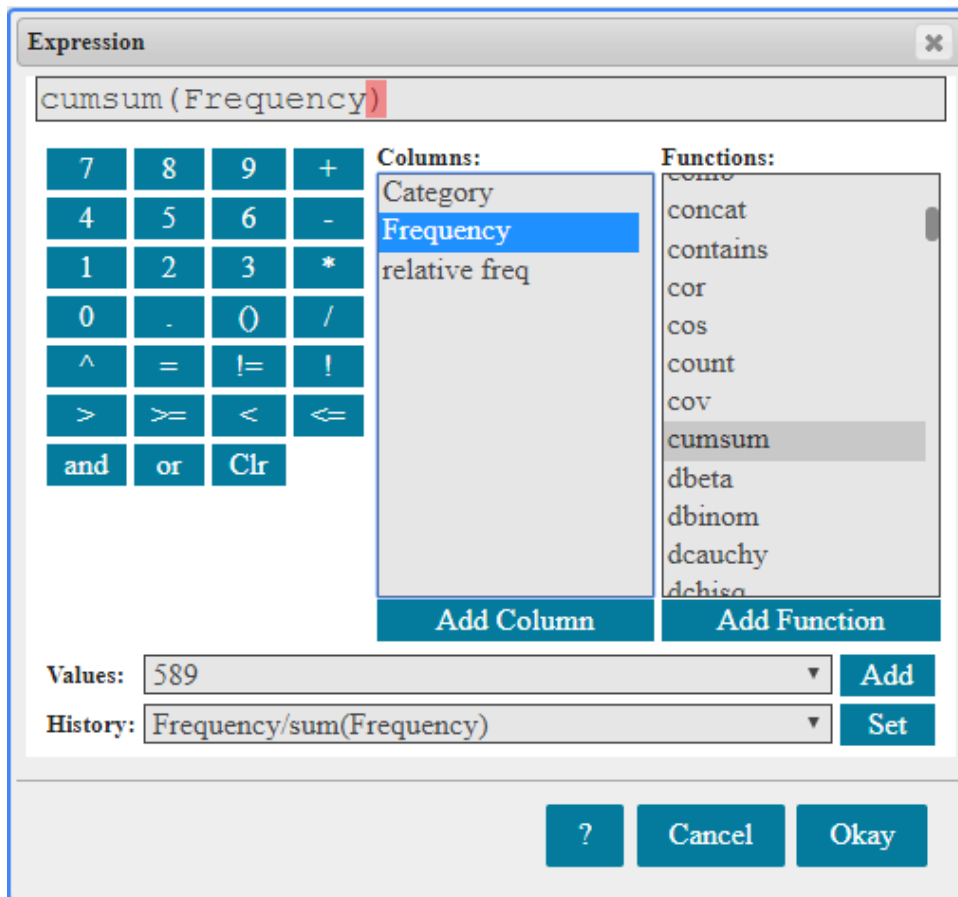
Figure 6: Cumulative Frequency Formula

Ok, so now we just click "Okay" and then "Compute." And we can see that a new column has been created with the cumulative frequency values for our data.

Figure 7: Cumulative Frequency Results

1.1.3 Relative Cumulative Frequency

Finally, let's find relative cumulative frequency for this dataset. This process is very similar to the process for finding regular cumulative frequency that we just did. First, start by clicking on "Data" then "Compute," "Multiple Expressions." This time I'll call our column "rel cum freq." Ok, now click on "Build" again. So remember that relative cumulative frequency simply takes the relative frequencies in each row and adds them up, totaling up to a final value of 1. This process requires the same "cumsum" function that we used before.
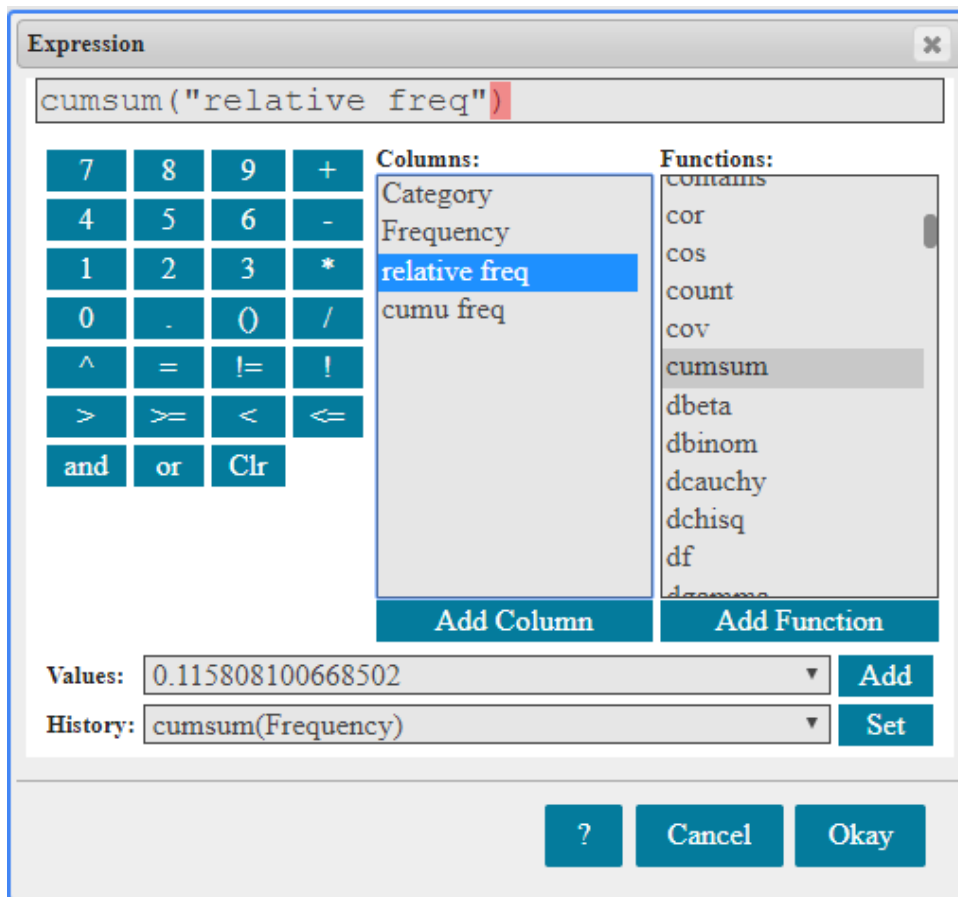
Figure 8: Relative Cumulative Frequency Formula

But this time, instead of using the "Frequency" column we will select the "relative frequency" variable instead. This tells StatCrunch to compute the cumulative sum of all of our relative frequency values, which is what we want. Now click on "Okay" and then on "Compute."

Row	Category	Frequency	relative freq	cumu freq	rel cum freq	var6
1	A	1131	0.22237515	1131	0.22237515	
2	B	1087	0.21372395	2218	0.4360991	
3	C	1362	0.26779394	3580	0.70389304	
4	D	589	0.1158081	4169	0.81970114	
5	E	917	0.18029886	5086	1	
6						

Figure 9: Relative Cumulative Frequency Results

Ok, and now after we click out (it'll sometimes popup these boxes and we can click out of those) we'll see our new column with the relative cumulative frequency values. Notice too that they add up to 1. This will always be true with relative cumulative frequency.

1.2 Raw Data

Okay, so now let's consider a second example. Say that instead of summary data, you were instead given a table of raw values and asked to find the same information.

Figure 10: Relative Cumulative Frequency Results

In this example, we have a column of data that contains letter grades for a certain class. We want to know the frequency, cumulative frequency, and relative cumulative frequency for each of these grades. In other words, we want to be able to answer questions like: How many A's, B's, and C's did we have? And if we chose a random student, what is the probability that their grade would be an A or a B? We need frequency counts and relative frequency in order to find the answers. Luckily, computing these values for raw data is much simpler than it is for summary data.
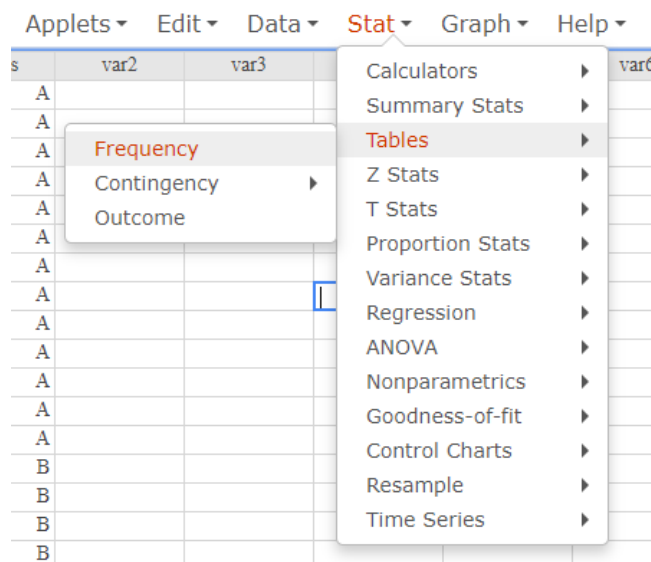

Figure 11: Select "Stats" , "Tables" , "Frequency"

So we just need to go to "Stats," and then "Tables," and then "Frequency." In the Frequency window, select the name of the column containing your data values. Here that's "Grades." Now under the "Statistics" box, we are going to click and drag to highlight all of the options. Then under "Output" click the box next to "Store in data table." Okay, now we have everything we need, and we can just hit "Compute."
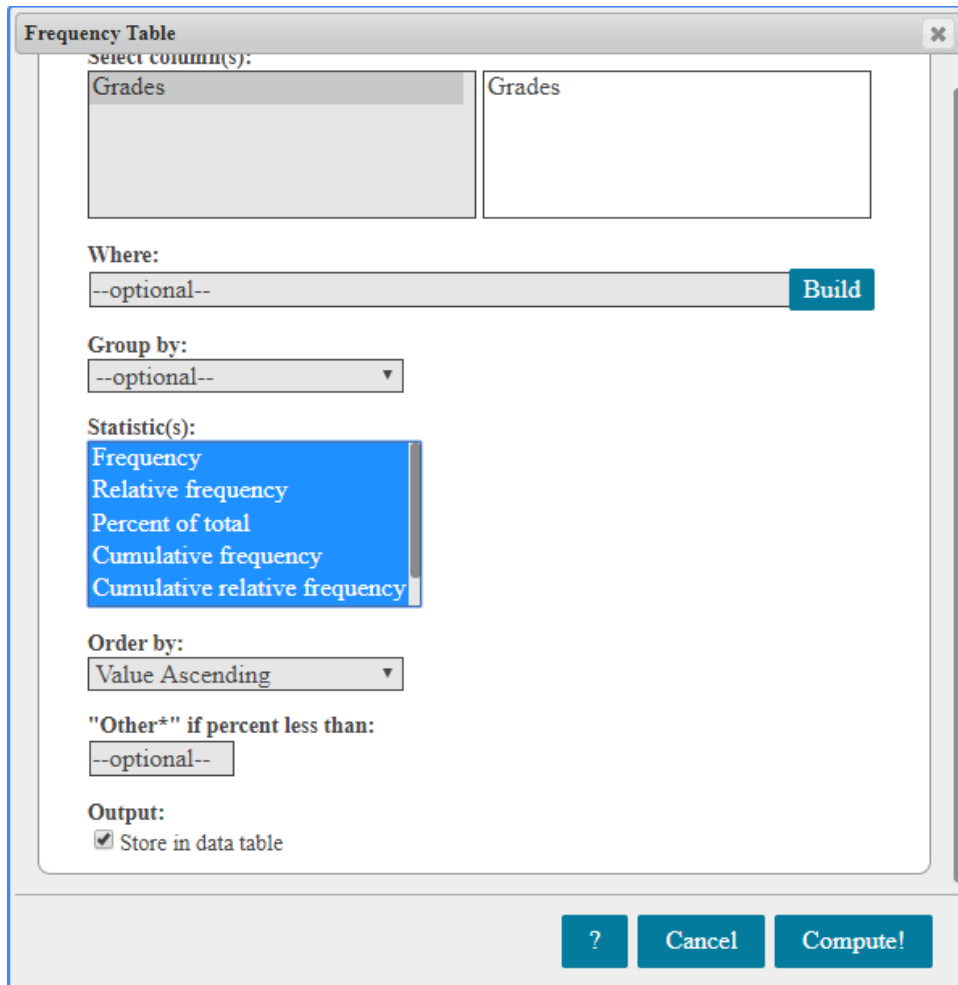
Figure 12: Frequency Window

Figure 13: Final Results

Ok, so now we can see that StatCrunch has created a bunch of new columns for us. We have one listing the Frequency values, one listing Cumulative Frequency, one listing the Relative Frequency values, and one listing Relative Cumulative Frequency values. And notice again that Relative Cumulative Frequency adds up to 1. And also, notice that the Cumulative Frequency values, here, total to 67. And that is the total number of grades that we had in our original data set.

Ok, so I hope you found this video helpful. If you are a UNT student, there will be some more links to other resources in the video description. Thank you for watching!

All figures produced using StatCrunch 2019, Pearson Education Inc.