

Tutorial 9: Splitting Ordinal/Continuous Variables (2 categories)

Description

This tutorial will convert a continuous or ordinal/integer variable into a categorical variable. This tutorial will specifically focus on splitting a variable into 2 categories.

In our dataset, we have variables that can be used as measured independent variables to understand their systematic influence on participants' responses. For example, in our project, we may want to see if participants' self-reported Liberal identity influences their response towards the person who made the offending comment.

The political identity variable (**pol_continuum**) was answered using a scale from 1-7 ranging from 1 = 'Strongly Conservative' to 7 = 'Strongly Liberal.'

Content

- 1. Find Desciptives
- 2. Create new variable
- 3. Check Descriptives Again

1. Find Desciptives

- a. Go to the 'Analyses' tab.
- b. Click on 'Explore' and then 'Descriptives'
- c. Move the **pol_continuum** variable into the variable window on the right and select the following statistics: Mean, Median, Standard Deviation.
- d. Based on the mean, median, and standard deviation, we can see from the descriptives that data is highly skewed.
- e. Examining the frequency distribution gives us an even clearer idea of how participants responded:

🚍 Data Analyses Edit						:
$ \begin{array}{c} & & & \\ \hline & & & \\ \hline & & & \\ \hline \\ \hline$	requencies Factor					ېلې Modules
Descriptives	Variables Variables Split by	Descriptives Descriptives N 180 Missing 0 Mean 5.60 Median 6.00 Standard deviation 1.33 Frequencies				
🗹 Frequency tables 🐣 📊		Frequencies of pol_continuum	0	0/ -1 T-1-1	Quere lating N	
✓ Statistics		Levels	counts	% of lotal	cumulative %	
Sample Size N Missing Percentile Values	Central Tendency C Mean Median Mode Sum	Suffictive Conservative Slightly Conservative Neither Conservative nor Liberal Slightly Liberal Somewhat Liberal Strongly Liberal	9 37 23 49 60	5.0% 20.6% 12.8% 27.2% 33.3%	6.1% 6.1% 26.7% 39.4% 66.7% 100.0%	
Dispersion Std. deviation Minimum	Distribution Skewness					
Variance Maximum Range S. E. Mean	Kurtosis Normality	References				

- f. Based on the distribution, we decided that it would be helpful to re-categorize this variable as:
 - i. People who are moderately Liberal (5 and 6; N =72)
 - ii. People who report strongly Liberal (7; N = 60)

2. Create new variable

- a. We can create these new categories using the IF' function-, meaning that IF the first logical expression is TRUE, insert a first value, but IF a second logical expression is TRUE, insert a second value, but IF a third logical expression is TRUE, then insert a third value.
 - i. **Note:** use == for 'equals' statement.
 - ii. You must close with three left parentheses (see formula below) because there are 3 commands otherwise you will get an error message.
- b. Scroll to end of spreadsheet and create new variable
- c. Name new variable and give a brief description
- d. In function window- specify code

=IF(pol_continuum == 5 or pol_continuum == 6, 1, IF(pol_continuum == 7, 2, IF(pol_continuum < 5, 0)))

i. This is basically saying that the new variable will be coded as 1 when a participant is moderately liberal and 2 when strongly liberal. We gave anyone less than a 5 a '0'.

=		Data Analy	/ses	Edit					:
Pas	ste C	➡ Ċ Edit	Setup	Compute	Transform Variables	Add	• Filter	Add • s E Delete Rows	, 9
							COMPUTED	VARIABLE	
							Newliber	al	
							moderateOto	extreme1	
							moderateote		
							Formula	£ - 6, 1	F(pol_continuum == 5 or pol_continuum == 1. IF(pol_continuum == 7. 2. IF(
								pol_	L_continuum < 5, 0)))
					<u></u>				
									Retain unused levels 🔘
NewR	elig	A NewEthni.!	🚽 Newli	beral®					Ectimated Marginal Means
1	1	_		1					Estimated Marginal Means
2	1			0					Apology cond
3	1			0					
4	0			2					
5	0			2					3.00 -
6	0			1					т –
7	1			2					
8	1			0					2.75
9	1			1					6.17 Ø
10	1			1					
12	0			2					ы ,
13	1			0					¥ 2.50 °
14	0			0					
15	0			1					
Ready	· 7	Filters 0	Row	count 180	Filtered 0 De	eleted 0 A	dded 61 Cel	s edited 10455	2.25 - 🔟

e. You can see that the recategorization worked when you run the frequencies on your new variable and get the same values for each category in the first frequencies you ran on the original Liberal variable.

Image: Second Secon	☰ Data Analyses Edit			1
Descriptives if Mr 432 if Konstription if Konstription if Mewitchnicity Split by NewKitchnicity if Maising Prequency tables if Missing Percentiles 25,50,75 Dispression Distribution Cut points for if equal groups Sum Dispression Distribution Cit deviation Minium Cit deviation Minium Station Maximum Cit deviation Maximum Cit deviation Maximum Cit deviation Minium Station Normality OR	Exploration T-Tests ANOVA Regression Freq	uencies Factor		Modules
a activits_plit a vertivits_plit a vertivits_plit a vertivits_plit split by spl	Descriptives	$\overline{\mathbf{G}}$	Strongly Liberal 60 33.3 % 100.0 %	
Sample Size Central Tendency Ø N Ø Missing Mean Percentile Values Median Out points for 4 equal groups Mode Percentiles 25,60,75 Sum Pispersion Distribution Std. deviation Minimum Skewness Variance Maximum Skewness Variance Maximum Skewness Narimum Skewness	di activist_split di NewReligion di NewEthnicity ✓ Frequency tables & di	Spiit by	Descriptives Newliberal N 180 Missing 0	
Dispersion Distribution Std. deviation Minimum Skewness Variance Maximum Std. deviation Minimum Skewness	✓ Statistics Sample Size N ♥ Missing	Central Tendency Mean	Frequencies	
Std. deviation Minimum Skewness Variance Maximum Kurtosis Range S. E. Mean Normality IQR Shapiro-Wilk	Cut points for 4 equal groups Percentiles 25,50,75 Dispersion	Mode Sum Distribution	Levels Counts % of Total Cumulative % 0 48 26.7 % 26.7 % 1 72 40.0 % 66.7 % 2 60 33.3 % 100.0 %	
	Std. deviation Minimum Variance Maximum Range S. E. Mean	Skewness Kurtosis Normality Shapiro-Wilk		

f. Filtering out the 'Not Liberals'

=	Data	a Analyses	Edit					:
Pas	ste	⊖ Edit	Compute	Transform	Add - Til	ters Rows		
						TERS	active C X active X active X active X active X actives X	⑦
	Filter 1	A.ID	A gender	🐣 age	🐣 religion	This is a data va	riable.	22.3% 100.0%
1	~	1	Female	18	Christian (in	-95	an angly maaran	33.5 % 100.0 %
2	2	2	Female	19	Muslim	-95		
3	1	3	Male	20	Christian (in	-99		
4	~	4	Female	19	No Religion	-95		
5	~	5	Male	24	No Religion	-95		
6	1	6	Female	17	No Religion	-99	Descriptives	
7	×	7	Female	19	Christian (in	-95		
8	1	8	Female	18	Muslim	-95	Descriptives	
9	1	9	Female	18	Jewish	-95	Newliberal	
10	×	10	Female	19	Christian (in	-95		
11	×	11	Male	19	No Religion	-95	N 180	
12	×	12	Female	19	No Religion	-95	Missing	
13	~	13	Male	18	Christian (in	-95		
14	1	14	Female	21	No Religion	-95	Francisco	
15 Readu	V ON	Filters 1	Row count 180	18 Filtered 0 Delete	NO Religion	-9§ Cells edited 10455	Frequencies	

=	E Di	ata Anal	yses	Edit												
Pa	iste	⊖ C Edit	Setu	up Compute	Transform Variables		Add • Delete	Filters	Add - Delete Rows							
							ROW F	ILTERS								
							+	Filter 1 $f_z = 1$ only n	1 Newliberal> noderate and	=1 strong liberals		active (• × +	>		
	Filter 1	🔒 ID		🐣 gender	🐣 age		🐣 religion		religion_7							
1	× .		1	Female		18	Christian (ir	n	-99							
2	×		2	Female		19	Muslim									
3	×		3	Male			Christian (ir			Descriptiv	ves					
4	× .		4	Female		19	No Religion		-99							
5	× .		5	Male		24	No Religion		-99	Descriptives						
6	× .		6	Female		17	No Religion		-99	Descriptives	uliberel					
7	~		7	Female		19	Christian (ir	1	-98		wilberal					
8	×		8	Female		18	Muslim		-98	N	132					
9	× .		9	Female		18	Jewish		-98	Missing	0					
10	× .		10	Female		19	Christian (ir	1	-99							
11	× .		11	Male		19	No Religion		-98							
12	1		12	⊢emale		19	No Religion		-99	Frequencie	S					
13	×	_		Male		18			-91							
14	×		14	Female		21	No Religion		-99	Frequencie	s of Newlil	peral				
15		Ellerer 4	15	⊢emale	11	18	No Religion	0.1	-99	Levels	Counts	% of Total	Cumulative %	-		
Read	iy 🍸 🤨	Filters 1	F	Row count 180	-iltered 48 Del	leted	0 Added 61	Cells e	dited 10455					-		

3. Check Descriptives Again

a. If you run frequency on NewLiberal- it should verify that you did this correctly

≡ Data Analyses Edit			:
Exploration T-Tests ANOVA Regression Fr	requencies Factor		Modules
Descriptives	Variables	Descriptives	
al MF032 al k_contempt purity_comp al activist_split al NewReligion	Split by	Descriptives Newliberal N 132 Missing 0	
Frequency tables Statistics	<u> </u>	Frequencies Frequencies of Newliberal	
Sample Size	Central Tendency	Levels Counts % of Total Cumulative % 1 72 54.5% 54.5%	
🗹 N 🗹 Missing	Mean	2 60 45.5% 100.0%	
Percentile Values	Median		
Cut points for 4 equal groups	Mode	>	
Percentiles 25,50,75	Sum		
Dispersion	Distribution		
Std. deviation Minimum	Skewness		
Variance Maximum	Kurtosis	References	
Range S. E. Mean	Normality	[1] The jamovi project (2021). <i>jamovi</i> . (Version 1.6) [Computer Software]. Retrieved from https://www.ja	movi.org.

-----END TUTORIAL-----

This Jamovi tutorial is a companion to a video tutorial and these materials were developed by:

Aline Hitti, Ph.D., University of San Francisco Saera Khan, Ph.D., University of San Francisco Sage Stefanick, University of San Francisco Erica Divinagracia, University of San Francisco Adolfo Barrales, University of San Francisco

This tutorial was made possible by an Open Education Resource grant awarded to the first two authors by Gleeson Library, University of San Francisco.



CHANGE THE WORLD FROM HERE