

Tutorial 15: Exploring Data - Skewness & Normality

Description

This tutorial will describe how you can explore your data to assess whether variables that have interval or ratio (ordinal/integer or continuous) measurement types are normally distributed or skewed. Sometimes skewness can be assessed using distribution graphs, as the one described in tutorial 14 but to get a more accurate assessment of skewness and normality it is better to make use of the 'skewness' statistic and the 'Shapiro-Wilk' test.

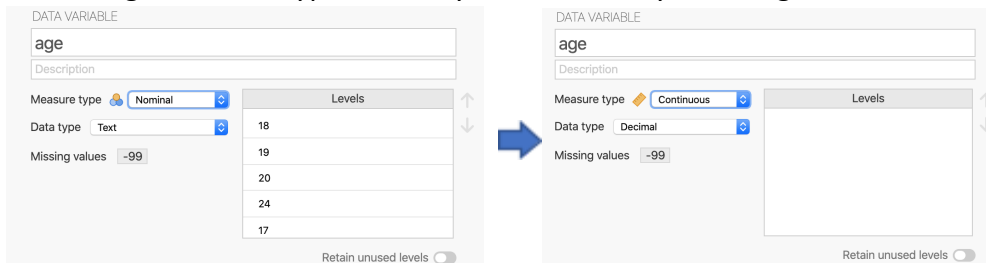
Note. We will work with the continuous variable *age* and the ordinal/integer variable *activism* in our dataset.

Content

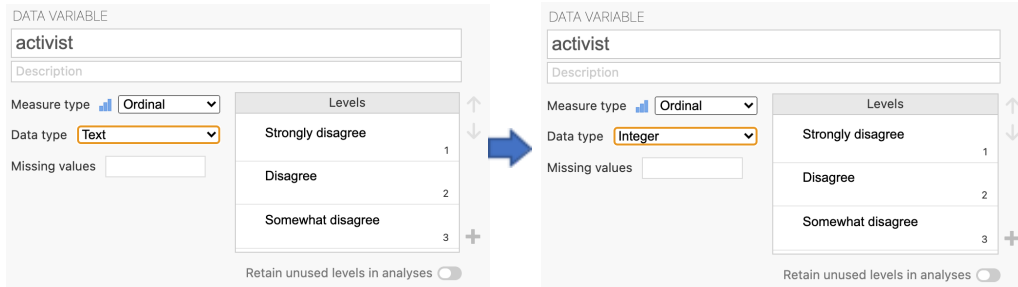
1. Check measure type for each variable
2. Exploring normality and skewness.

Check measure type for each variable

1. Select the '*age*' variable by double clicking on the column header with the *age* variable.
2. Change the Measure Type to 'Continuous'. This will automatically change the Data Type to 'Decimal' but check that this happens, if it does not change automatically, you will have to change the Data Type manually to 'Decimal' by selecting it.



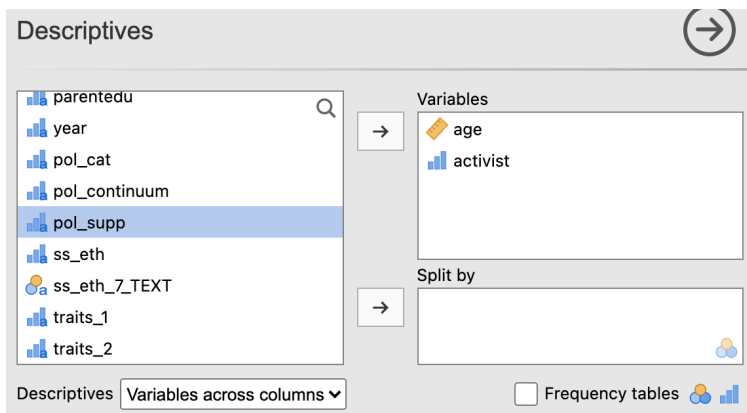
3. Now, select the '*activism*' variable by double clicking on the column header with the *activism* variable.
4. Change the Measure Type to 'Ordinal'. Then change Data Type to 'Integer'.



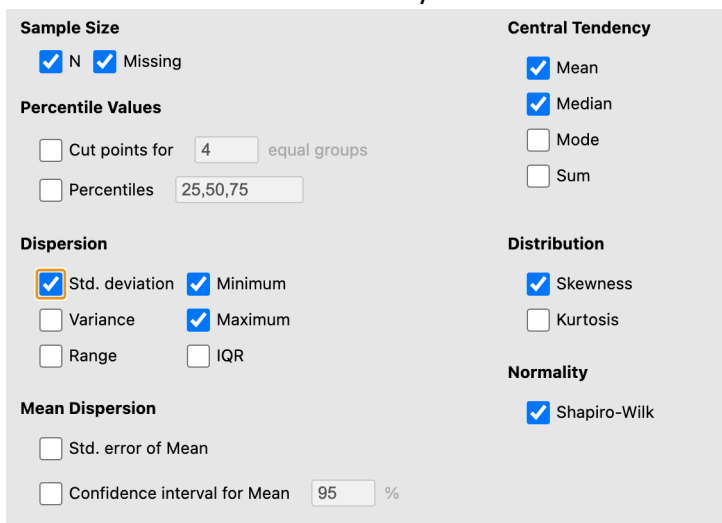
Exploring normality and skewness.

We will obtain descriptive statistics that are specific to skewness and normality through the 'Explore' analyses button.

1. To obtain descriptives select the 'Analyses' tab in the top left corner. Under this tab also in the top left corner click on "Explore" and select 'Descriptives'.
2. Move the following variables (*age*, *activist*) to the variable window by selecting the variables in the window on the left and clicking the arrow for the window on the right labeled "Variables"



3. Click on the 'Statistics' tab and under the 'Distribution' section and check 'Skewness'
4. Then under the 'Normality' section and check 'Shapiro-Wilk'.

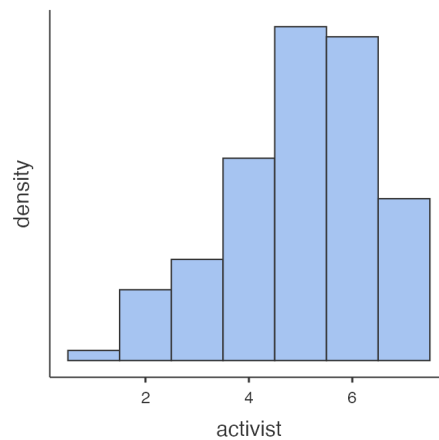
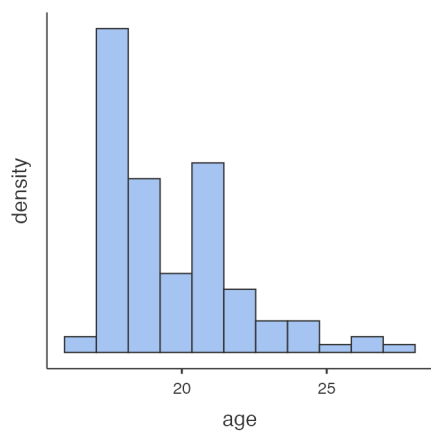


5. Interpreting the output requires that you interpret the values reported for skewness and the Shapiro-Wilk test of normality. See the table below which appears in the right-side output window in Jamovi. Note we are looking at the numbers for “Skewness” and “Shapiro-Wilk p”:

Descriptives

Descriptives		
	age	activist
N	180	180
Missing	0	0
Mean	19.7	4.92
Median	19.0	5.00
Standard deviation	1.99	1.46
Minimum	17.0	1
Maximum	28.0	7
Skewness	1.59	-0.545
Std. error skewness	0.181	0.181
Shapiro-Wilk W	0.830	0.926
Shapiro-Wilk p	<.001	<.001

6. Skewness values for a normal distribution are close to zero, if the skewness is farther away from 0 (either in the positive or negative direction) then you can conclude that the distribution is skewed. Thus skewness for **age** is at 1.59 (positively skewed), and skewness for **activist** is at -0.545 (negatively skewed). See graphs below to reference shape of the distributions.
7. However, the Shapiro-Wilk test will tell you whether the distribution is statistically significantly non-normal. Therefore, p-values for this test that are less than .05 indicate that the distribution deviates from normality. “Shapiro-Wilk p” indicates the p-value for the test and tells us whether the distribution is statistically significantly different from normal and in this case since the p-values are less than 0.05 (<.001) for both variables then the distributions for both variables are not normal.
8. When normality is violated then the variables must be transformed in order to be used in analyses. Some transformations are covered in other tutorials such as transforming the variable to a categorical measurement type. Other possible transformations include log transformations which keep the measurement as ratio/continuous but are beyond the scope of these tutorials.

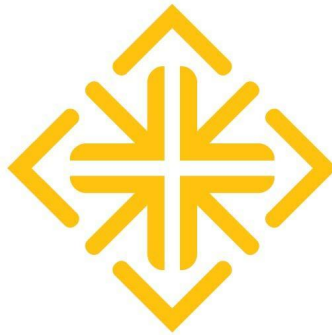


-----END TUTORIAL-----

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.



**UNIVERSITY OF
SAN FRANCISCO**

CHANGE THE WORLD FROM HERE