Al-Mustaqbal University
College of Health and Medical Technologies
Lecturer Mohammed Qasim Al Azzawi
mohammed.qasim.obayes@uomus.edu.iq

Frequencies and Distributions

This is the most important univariate procedure. Conducting it properly, and interpreting the output rigorously, will enable you to understand the major attributes of the frequency distribution of each variable.

1. The distribution of variables - histograms and frequency statistics

GO TO Analyze/ Descriptive Statistics/ Frequencies. Select the variables and move them into the box "Variable(s) as shown in Figure 1." Unless the variables you have chosen are categorical or dummy (i.e. - they have only a few discrete possible values), deselect the option "Display Frequency Tables." Otherwise, you will generate too many pages of output.

Figure (1): Frequencies Interface

Note: Conduct the frequencies procedure twice - Once for continuous variables (deselecting the option "Display Frequency Tables") and once for categorical and dummy variables (this time choosing the option "Display Frequency Tables"). Now you must instruct SPSS to construct a histogram for each of the chosen variables. Click on the button "Charts." Choose to draw a histogram with a normal curve - select the option "Histogram" and click on the box to the left of the title "With normal curve. " Click on "Continue."

Al-Mustaqbal University
College of Health and Medical Technologies
Lecturer Mohammed Qasim Al Azzawi
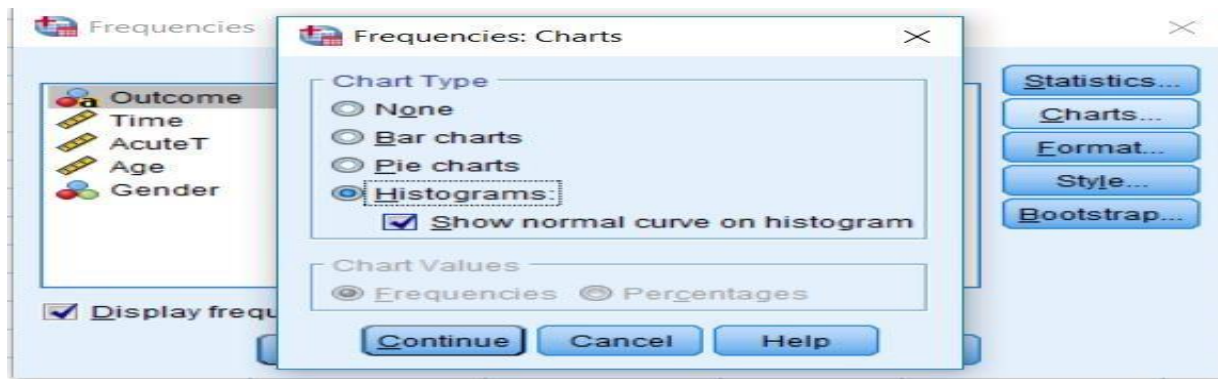mohammed.qasim.obayes@uomus.edu.iq

Figure (2): Selecting type of chart

What is the normal distribution and why is it so important?
 A variable with a normal distribution has the same mode, mean, and median, i.e. - its most often occurring value equals the average of values and the mid-point of the values.
Visually, a normal distribution is bell-shaped (as shown in Figure 3.) the "idealized normal curve- the left half is a mirror image of the right half. The importance stems from the assumption that "if a variable can be assumed to be distributed normally, then several inferences can be drawn easily and, more importantly, standardized tests (like the T and F tests) can be applied." In simpler terms: "normality permits the drawing of reliable conclusions from statistical estimates.
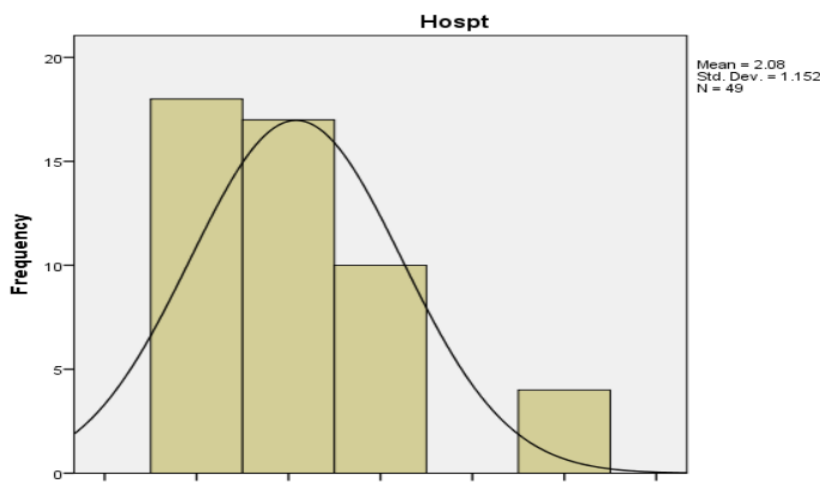

Figure (3): Histogram with normal distribution.

Al-Mustaqbal University
College of Health and Medical Technologies
Lecturer Mohammed Qasim Al Azzawi
mohammed.qasim.obayes@uomus.edu.iq

We also want to obtain **descriptives**. Click on the button "Statistics." As in Figure 4. Select the options as shown in Figure 5. These statistics cover the list of "descriptive statistics.

The options under "Percentile values" can assist in learning about the spread of the variable across its range.

Figure (4): Applying descriptives by selecting statistics

- Click on "Continue."
- Click on OK.
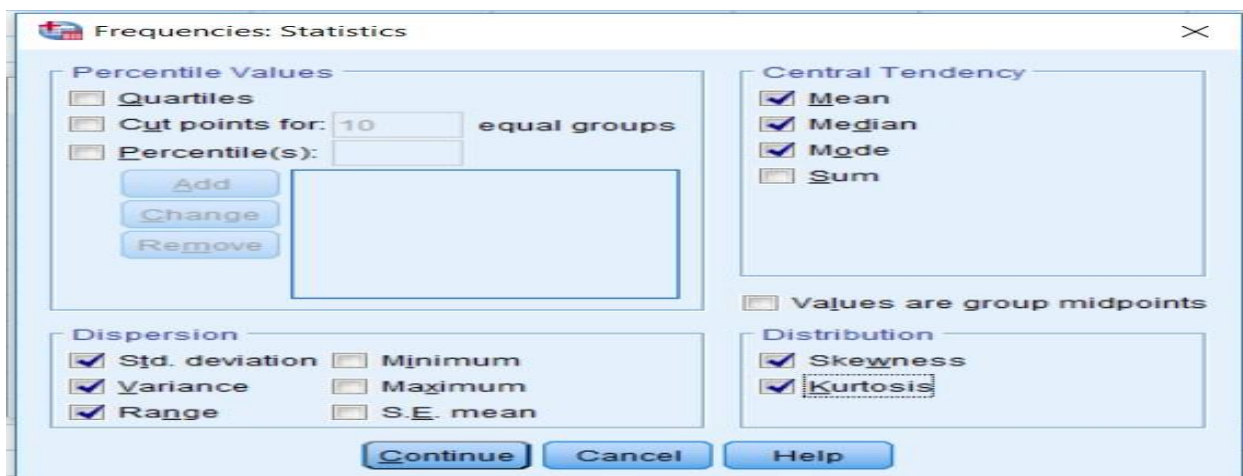- Click on "Continue."
- Click on OK.

Figure (5): selecting required statistics

The output will have one frequency table for all the variables and statistics chosen and one histogram for each variable.

Al-Mustaqbal University
College of Health and Medical Technologies
Lecturer Mohammed Qasim Al Azzawi
mohammed.qasim.obayes@uomus.edu.iq

In the next three graphs, the heights of the bars give the relative frequencies of the values of variables. Compare the bars (as a group) with the normal curve (drawn as a bell-shaped line curve). All three variables seem to be left heavy relative to the relevant normal curves, i.e. - lower values are observed more often than higher values for each of the variables, see Figure 6.

On the lower-right corner, the chart provides the most important statistics - standard deviation, mean, and sample size. (The other statistics - like the median, mode, range, skewness, and kurtosis) are usually more visually identifiable from a histogram. The mode is the highest bar, the median has half the area (under the shaded bars) to its left, and the skewness and kurtosis are measures of attributes that are easily identifiable, see Figure 7.

→ Frequencies

Warnings
Treat is a string so a histogram cannot be produced.

Statistics

		Hospt	Treat
N	Valid	49	49
	Missing	0	0
Mean		2.08	
Median		2.00	
Mode		1	
Std. Deviation		1.152	
Variance		1.327	
Skewness		1.200	
Std. Error of Skewness		.340	
Kurtosis		1.122	
Std. Error of Kurtosis		.668	
Range		4	

Figure (6): output of the selected statistics

Frequency Table

Hospt

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	18	36.7	36.7	36.7
	2	17	34.7	34.7	71.4
	3	10	20.4	20.4	91.8
	5	4	8.2	8.2	100.0
	Total	49	100.0	100.0	

Treat

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Imipramine	18	36.7	36.7	36.7
	Lithium	18	36.7	36.7	73.5
	Placebo	13	26.5	26.5	100.0
	Total	49	100.0	100.0	

Figure (7): Frequency table of variables