People's Democratic Republic of Algeria
University Med Khider of Biskra
Faculty of SNVSTU

Practical Protocol 01

Dr. Ben Gherbal Hanane

Data Analysis in Biosciences — Level: L3 biology
Email: hanane.bengherbal@univ-biskra.dz

Objective

The objective of this protocol is:

1. To guide how to enter data into SPSS and calculate descriptive statistics (central tendency and dispersion).
2. To show how to draw a histogram with a superimposed Gaussian (normal) curve.
3. To interpret the kurtosis coefficient, which indicates whether the distribution is normal, flat, or peaked.

Example

You measure the **weight of 20 lab mice** (in grams):

Weight (g)	22.0	22.5	23.0	23.5	24.0	24.5	25.0	26.0	27.0	29.0	30.0	31.0
Frequency	8	6	5	4	3	2	1	2	1	1	1	1

1. Calculate the central tendency parameters of this sample.
2. Calculate the dispersion parameters of this sample.
3. Draw the histogram of this sample and superimpose the Gaussian curve.
4. Deduce the shape of the distribution.

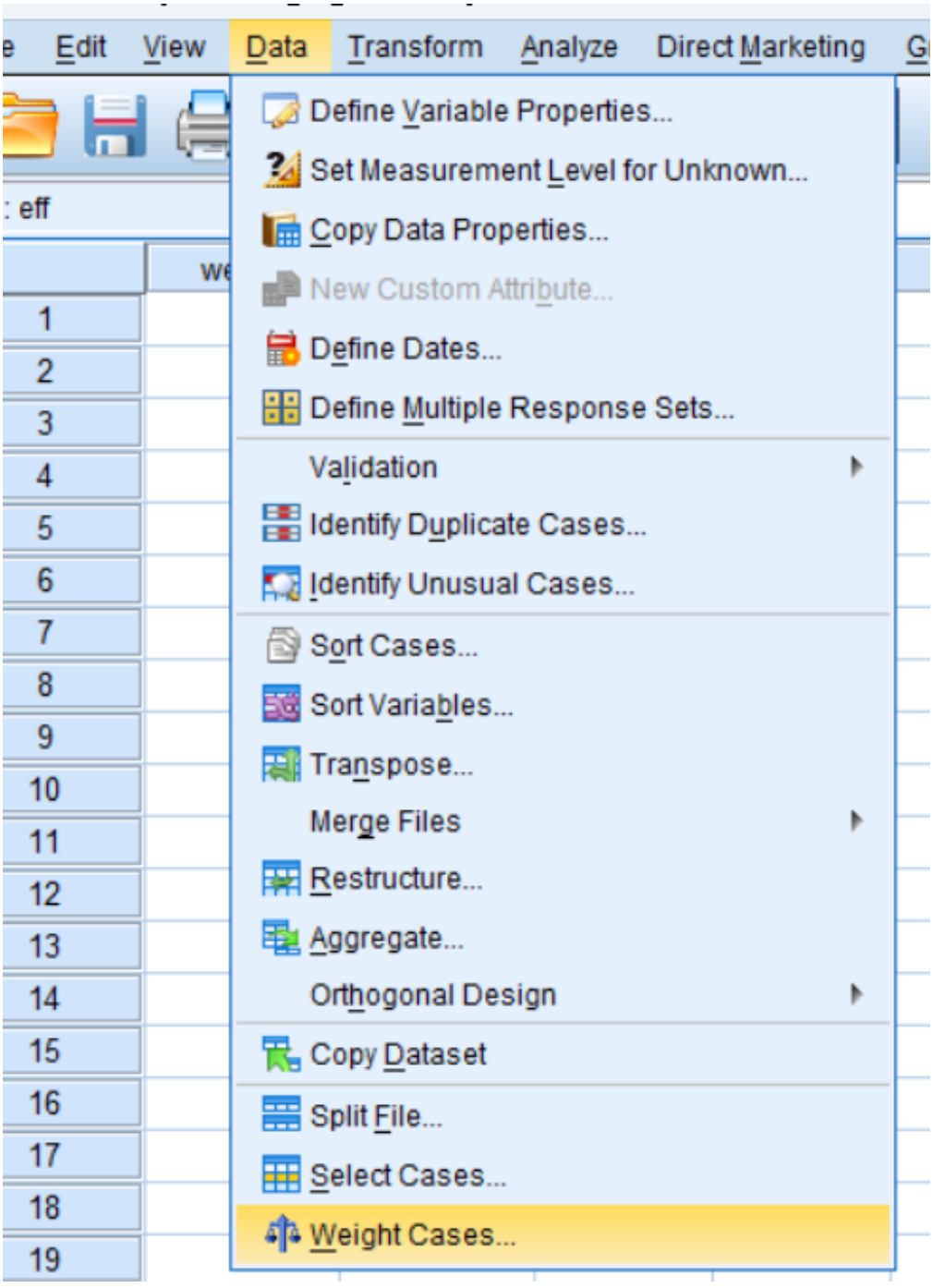
Solution

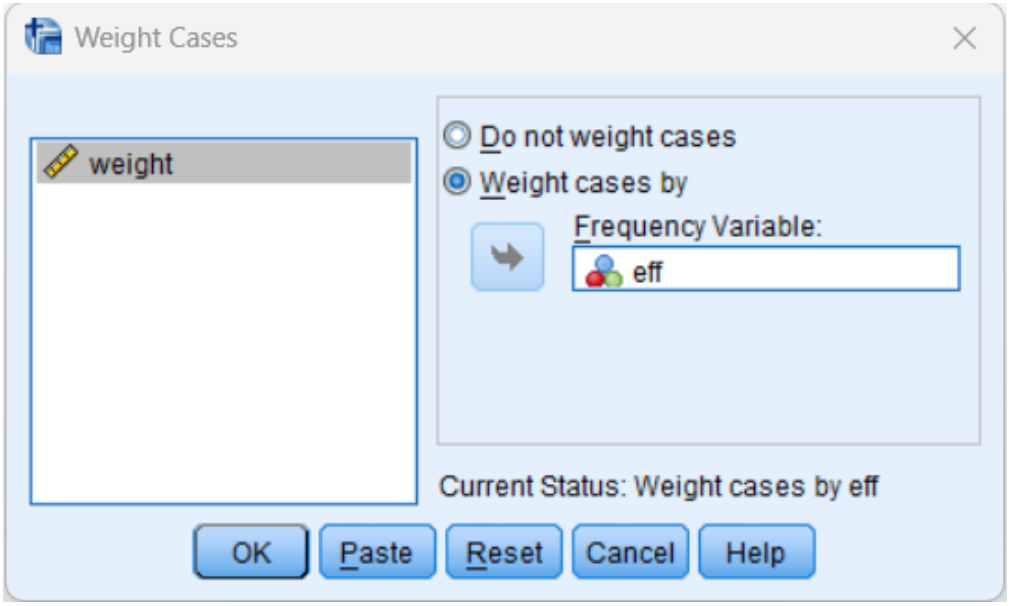
Step 1: Accessing descriptive statistics Launch SPSS and enter the data.

3 : eff

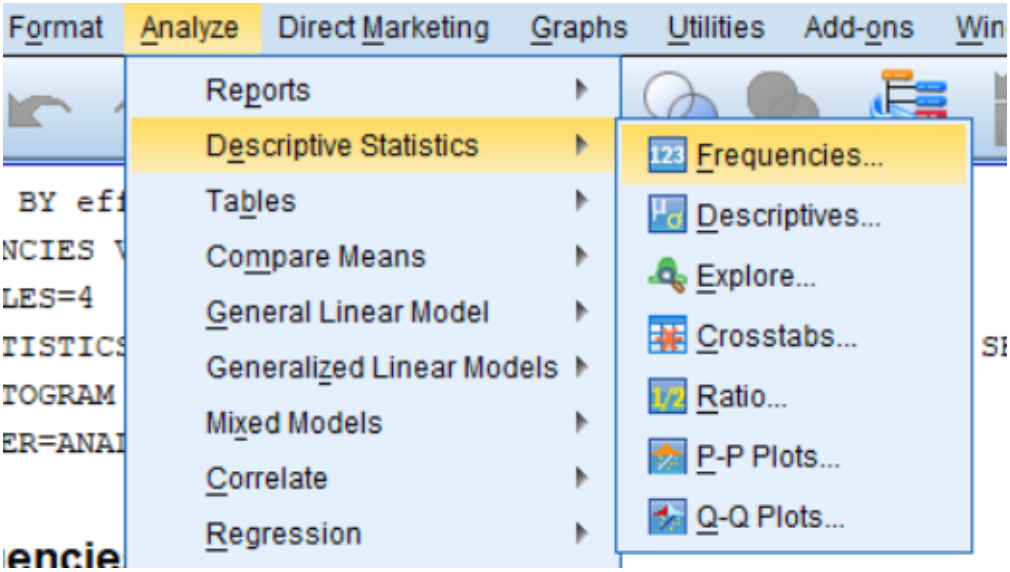
	weight	eff
1	22,00	8,00
2	22,50	6,00
3	23,00	5,00
4	23,50	4,00
5	24,00	3,00
6	24,50	2,00
7	25,00	1,00
8	26,00	2,00
9	27,00	1,00
10	29,00	1,00
11	30,00	1,00
12	31,00	1,00
13	.	.

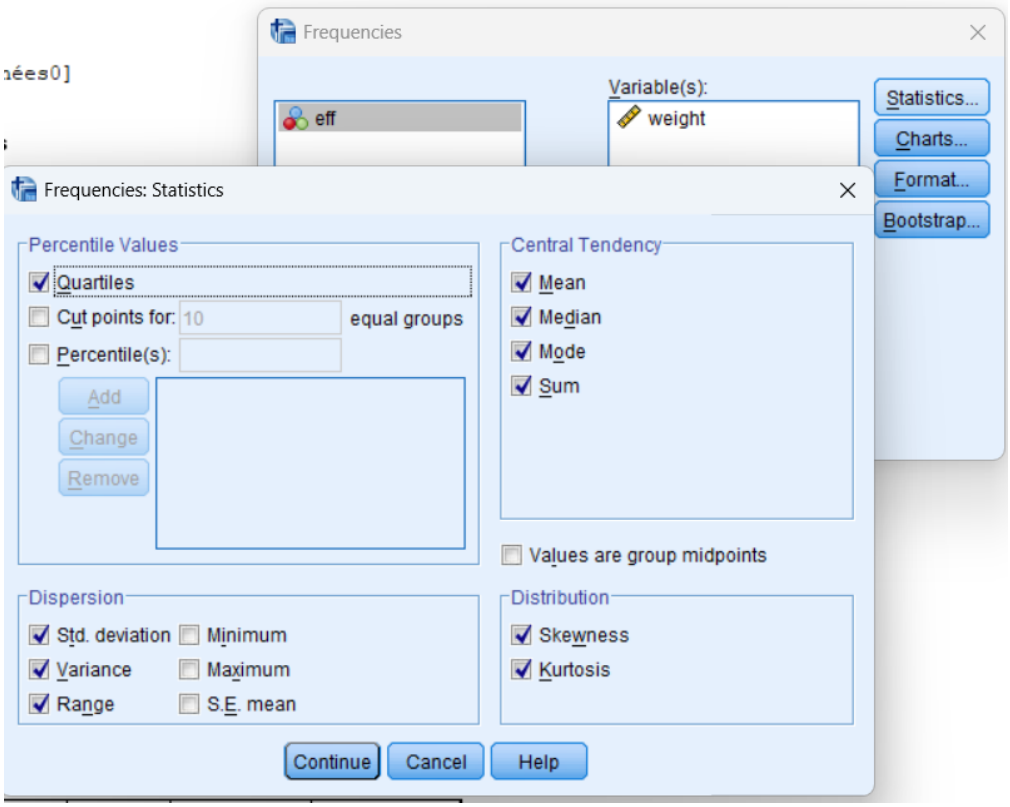
In the menu bar, choose Data → Weight Cases, then select Weight cases by frequency, and click OK


Edit
View
Data
Transform
Analyze
Direct Marketing
Define Variable Properties...
Set Measurement Level for Unknown...
Copy Data Properties...
New Custom Attribute...
Define Dates...
Define Multiple Response Sets...
Validation
Identify Duplicate Cases...
Identify Unusual Cases...
Sort Cases...
Sort Variables...
Transpose...
Merge Files
Restructure...
Aggregate...
Orthogonal Design
Copy Dataset
Split File...
Select Cases...
Weight Cases...

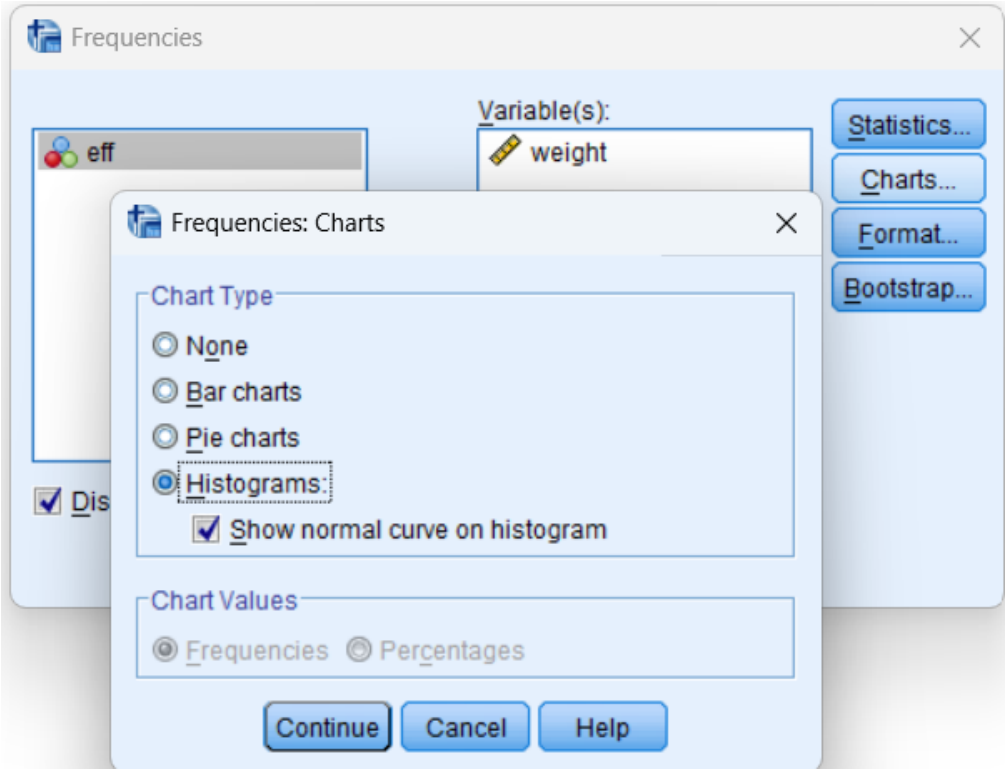


Choose *Analyze* → *Descriptive Statistics* → *Frequencies*. Select the variable (mice weight) and move it to the variable box. Click on *Statistics* and check the following: mean, median, mode, standard deviation, variance. Also check *Skewness* and *Kurtosis* to obtain information about the distribution shape.





Step 2: Creating a histogram with a Gaussian curve Go to *Analyze* → *Descriptive Statistics* → *Frequencies* → *Charts*. In the dialog box, select *Histogram* and check *Show normal curve on histogram*.



Results

Descriptive Statistics Table

Statistics

weight

N	Valid	35
	Missing	0
Mean		23,8571
Median		23,0000
Mode		22,00
Std. Deviation		2,29633
Variance		5,273
Skewness		1,843
Std. Error of Skewness		,398
Kurtosis		3,010
Std. Error of Kurtosis		,778
Range		9,00
Sum		835,00
Percentiles	25	22,5000
	50	23,0000
	75	24,5000

Gaussian Distribution and Shape Coefficients

The Gaussian curve, or normal distribution, is a symmetric function characterized by two main parameters: the mean and the standard deviation. However, when analyzing distributions that deviate from this ideal shape, two important concepts emerge: the coefficient of kurtosis and the coefficient of skewness.

Coefficient of skewness

The skewness coefficient measures the degree of asymmetry of a distribution around its mean. If the distribution is symmetric, like the Gaussian curve, this coefficient is zero.

1. **Zero skewness:** The distribution is perfectly symmetric (as in the Gaussian distribution).
2. **Positive skewness:** The distribution is skewed to the right.
3. **Negative skewness:** The distribution is skewed to the left.

Coefficient of kurtosis

The kurtosis is a parameter that measures the degree of flattening or sharpness of the distribution of a random variable.

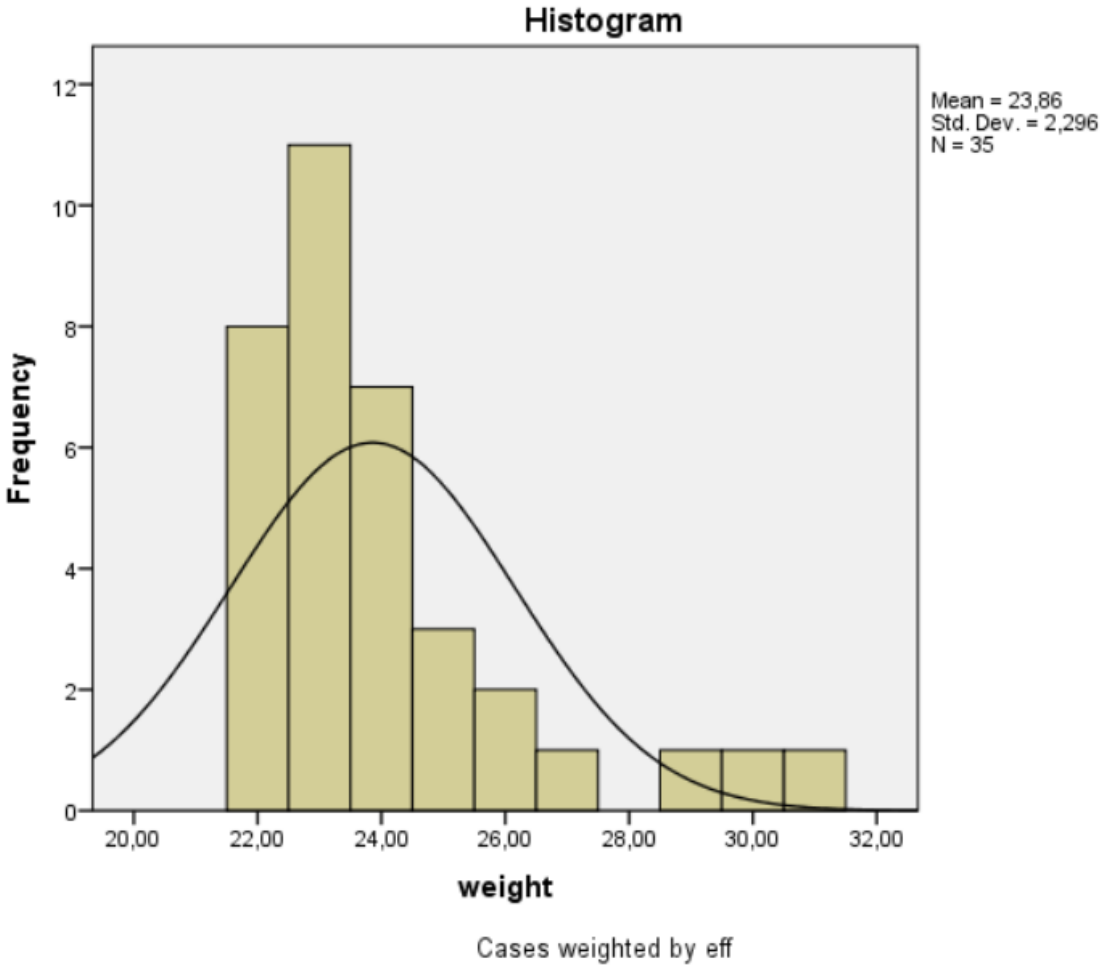
1. **Zero kurtosis:** This corresponds to the normal or Gaussian distribution.
2. **Negative kurtosis:** The distribution is flat (platykurtic).
3. **Positive kurtosis:** The distribution is sharply peaked (leptokurtic).

Conclusion

From the results:

- The skewness coefficient is slightly positive (≈ 1.843), so the distribution is slightly skewed to the right.
- The kurtosis coefficient is positive (≈ 3.01), so the distribution is sharply peaked.

Thus, the distribution deviates from normality: it is skewed to the right and more sharply peaked than a Gaussian distribution.


Histogram
Mean = 23,86
Std. Dev. = 2,296
N = 35
Frequency
0
2
4
6
8
10
12
20,00
22,00
24,00
26,00
28,00
30,00
32,00
weight
Cases weighted by eff