Exercise 2

In the file SAT scores.sav, the performance of some students in the grammar and math exams (verbal math, respectively) is recorded. Verify statistically the claim that students are equally good in both subjects.

Report: In this particular problem, we want to test whether the performance of students is equally good in grammar (verbal) and math. This is a problem of testing equality of means for two populations. Since our measurements are taken on the same experimental units for both subjects, we conclude that our samples are not independent. This experiment is of the before-after type. To answer our initial question, we calculate the differences between verbal and math scores, transforming our problem into a test of the population mean. To use the paired T-test, the following assumptions should be met for our sample:

1. It should be random.
2. It should not have extreme values in a percentage greater than 10%.
3. It should come from a population that is reasonably described by the normal distribution. From the boxplots, it is evident that there are three extreme observations in the sample values of the differences in student performance in grammar and math, with increasing numbers 124, 197, 130, and corresponding values 190, -270, -280, respectively. (see boxplots 1, 2, 3, 4)
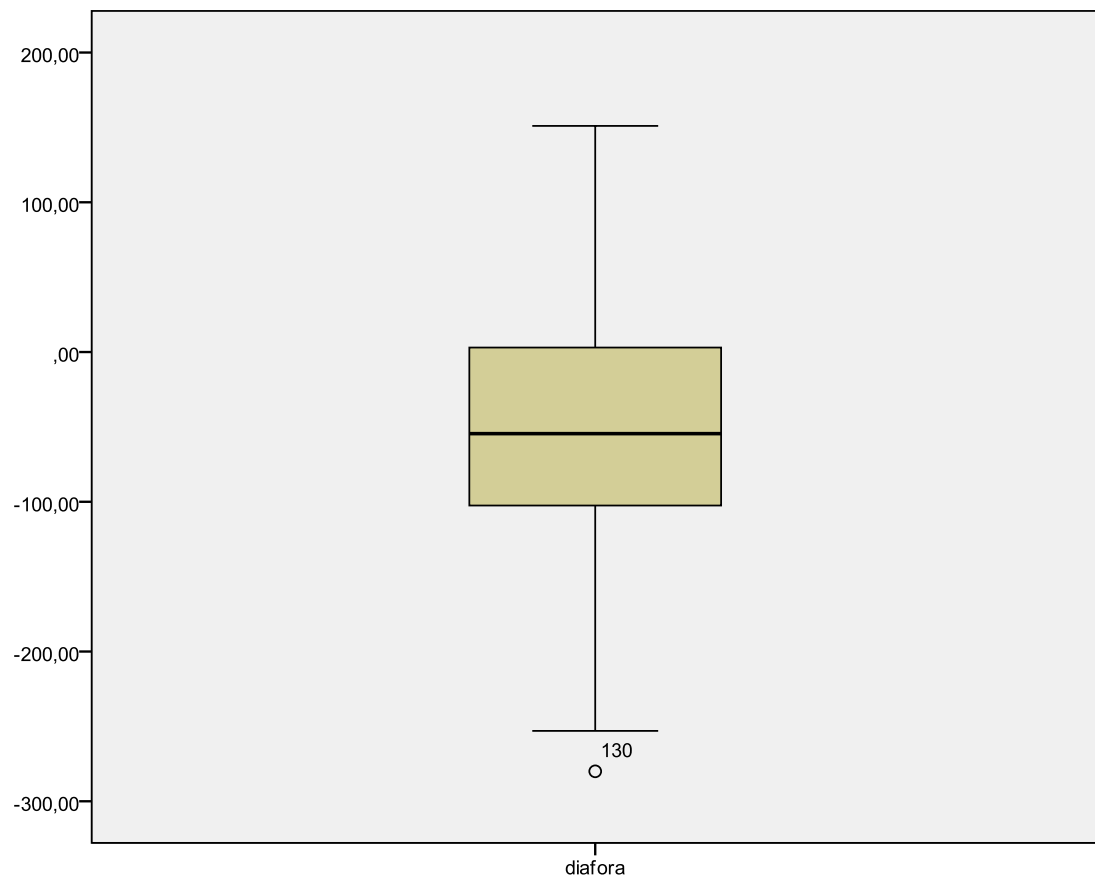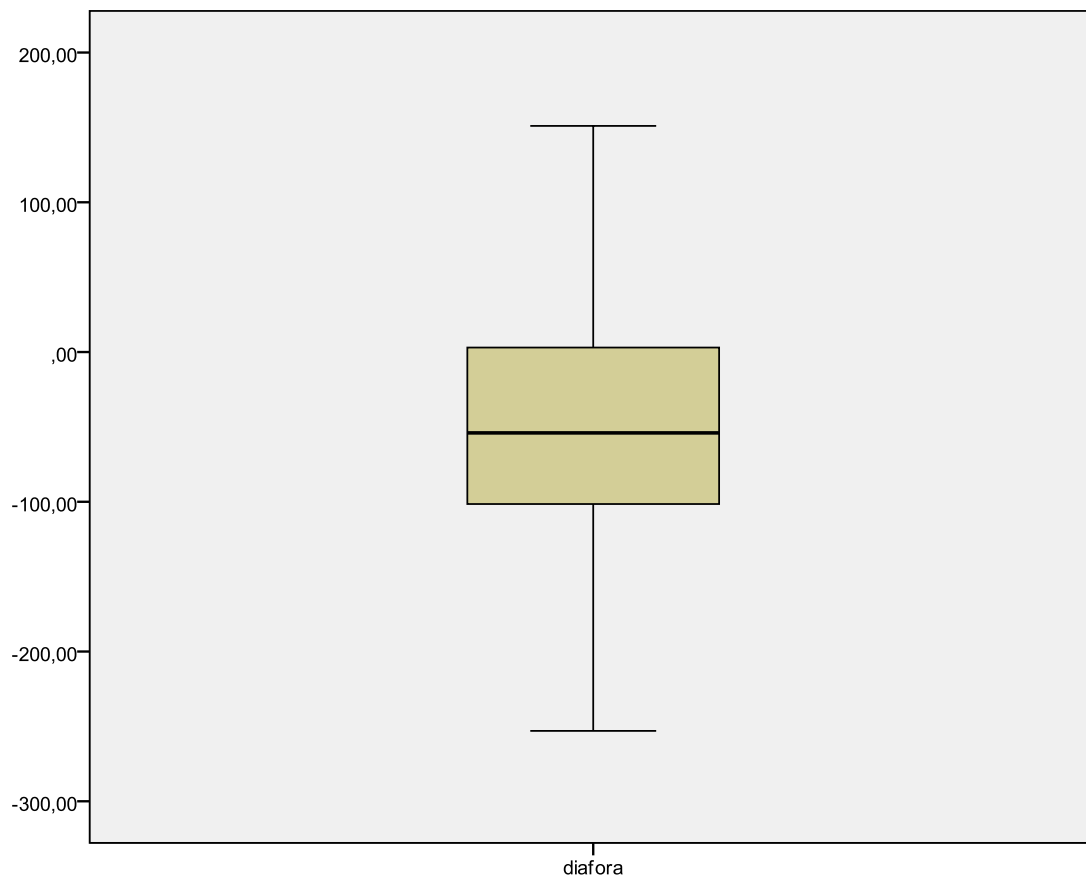
Boxplot 1

diafora
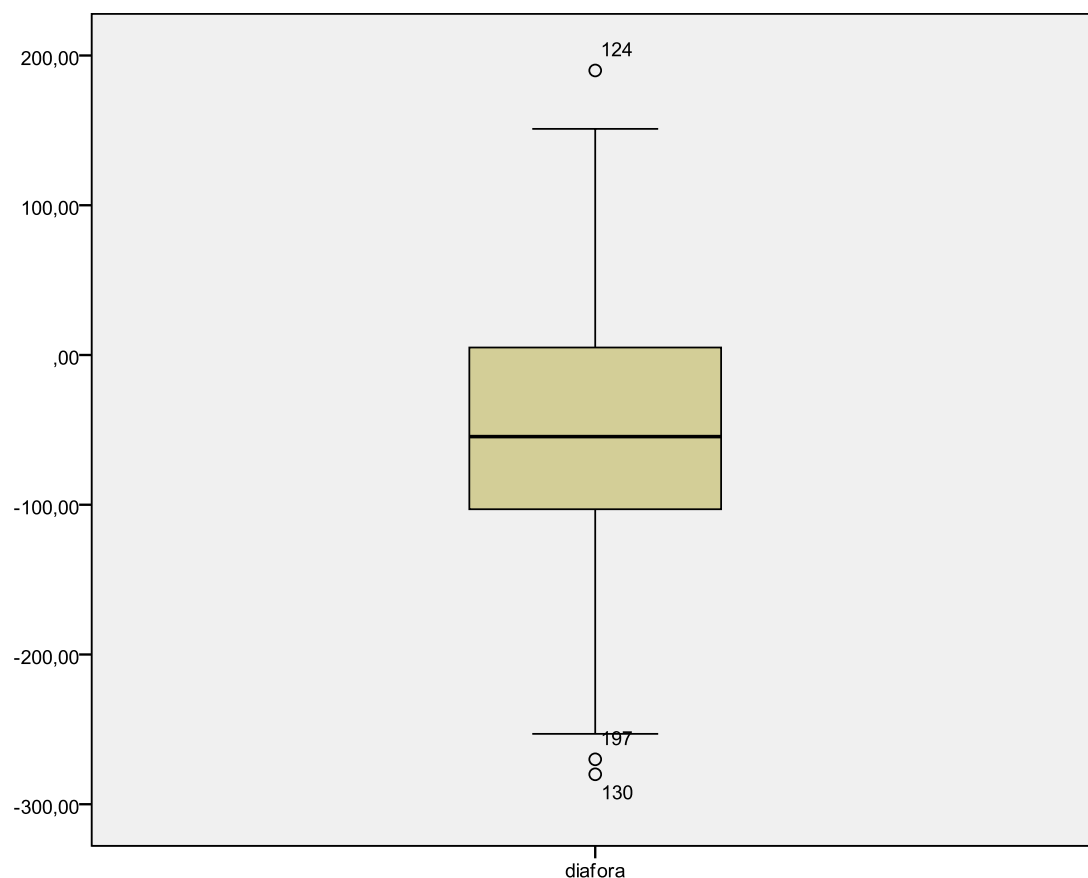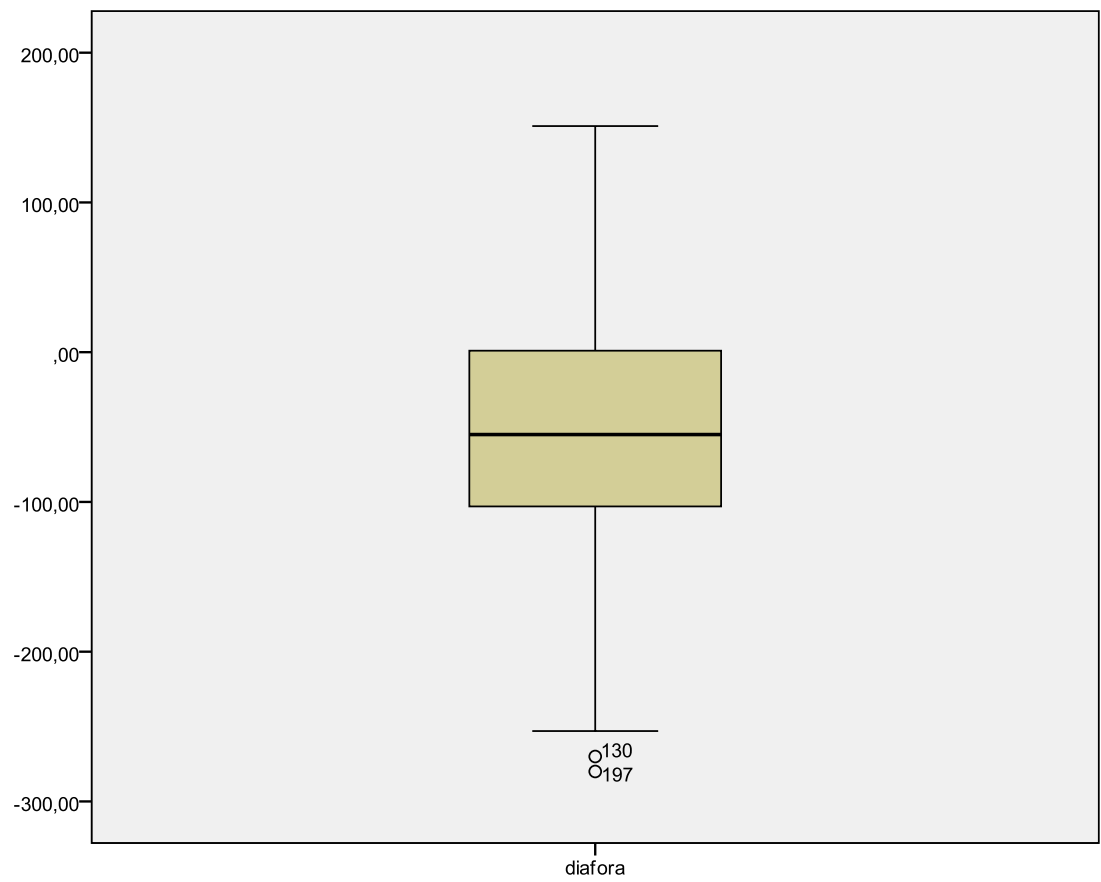
Boxplot 2

Boxplot 3

Boxplot 4

These observations are excluded from further analysis because their total number does not exceed 10% of the observations. Next, we check whether the 195 sample observations of the difference come from a normal population. The critical probability of the Shapiro-Wilk test is p = 0.320. This means that the hypothesis of a normal distribution cannot be rejected at a 5% significance level. Therefore, a parametric paired T-test will be used to test the hypothesis under study.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
diafora	,041	196	,200*	,992	196	,320

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

As a conclusion from the implementation of the paired T-test, it was found that the performance in grammar and mathematics differs statistically significantly (p < 0.001). Specifically, the mean performance in mathematics is 53.70769 points better than in grammar. A 99% confidence interval for the difference between verbal and math scores is (-64.97051, -42.44488).

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Verbal	595,5692	195	71,53954	5,12305
	Math	649,2769	195	65,09181	4,66132

		Paired Differences			95% Confidence Interval of the Difference		
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t
Pair 1	Verbal - Math	-53,70769	79,74406	5,71059	-64,97051	-42,44488	-9,405