Mention the way you could carry out a comparison between means of two samples (t-tests).

The first thing we need to mention is the issue that all the variables we are working with have a distribution that is not normal. We saw this in the activities of point 3. We even performed Box-Cox transformations with their corresponding lambdas. However, we still did not achieve normality. Therefore, we have to do non-parametric statistics. In other words, initially, we cannot perform Z-tests or Student's t-tests.

Now, to address one of the research questions, we can outline the procedure using a Student's t-test.

The question is: Do high levels of solar radiation affect ozone concentration in the air?

The procedure to answer this would be as follows:
1. Create samples. One sample would be of ozone data when there is low solar radiation. For example, we could use the samples taken during the night. The other sample would be of ozone data when there is high solar radiation. For example, taking ozone samples from 11 am to 4 pm.
2. Test for normality. Here, we would need to obtain samples with normality in these new samplings. If normality is not achieved, we could try Box-Cox transformations again. For the purposes of this exercise, we will assume that we do achieve normality (although we will later perform non-parametric statistical tests if our data still lack normality).
3. Test variances. This is to check if the variances between the samples with low and high solar radiation change. In the end, they are used to determine the degrees of freedom for the Student's t-test.
4. T-test. For this question, our null hypothesis would be that the mean ozone concentration is less than or equal when there is high solar radiation compared to low solar radiation. Thus, our alternative hypothesis would be that ozone concentration levels are higher when solar radiation is high. Therefore, we would conduct a one-tailed test, specifically a right-tailed test.