Problem n.2

The WWF is studying the quality of the waters of the Italian seas. To compare the Adriatic and Tyrrhenian seas, the WWF researchers perform identical and independent measurements on 8 protected sea areas of the Adriatic sea and on 10 protected sea areas of the Mediterranean sea. The file `waterquality.txt` contains the values of 7 pollution parameters measured by the researchers on the 18 sea areas. The higher the pollution parameter, the more polluted the water.

a) For each pollution parameter, perform a permutation one-sided test to look for possible statistical superiority of the quality of the waters of the Tyrrhenian sea. In detail, for each pollution parameter, use the difference of the sample means as test statistic and use 5000 random permutations with random seed equal to 2021 to estimate the permutational distribution. Report the value of the 7 test statistics and their corresponding p-values.

b) For which pollution parameters the waters of the Tyrrhenian sea can be considered superior to the waters of the Adriatic sea if the researchers want to limit the false discovery rate to a maximum value of 5%?

c) For which pollution parameters the waters of the Tyrrhenian sea can be considered superior to the waters of the Adriatic sea if the researchers want to impose a probability at most 5% that at least one of the non-superior purity parameter is judged as superior?

Upload your results here:
https://forms.office.com/Pages/ResponsePage.aspx?id=K3EXCvNtXUKAjjCd8ope612LHtvIHvFEsEi2L6mhPg
lUQ0lUTzZTTUZPROhKQkFWSjRMSU9RRFEOQi4u