Seminar 11
Non-Parametric Testing, Rank Tests

1. We monitor salaries in a hypothetical company. In the `R-script` for this seminar You can find the data sample containing the salaries in 2023 and 2022 for 40 different employees.

   (a) Firstly, check if the salaries in 2023 can be assumed to be **normally distributed**. Are there any **outliers**?

   (b) Test the hypothesis that the median salary in 2023 is equal to 46 against the alternative that it is smaller. Based on the exploratory data analysis, choose appropriate test. Use both step-by-step construction of the test statistic as well as a built-in function in R. HINT: Use the **sign test**. Can you use the Wilcoxon test in this example?

   (c) Test the same hypothesis for the value 45. Why has the result changed?

   (d) Test the hypothesis that the median salary in 2023 is equal to the median salary in 2022 against the **right-sided** alternative. Based on the exploratory data analysis, choose appropriate test. Use both step-by-step construction of the test statistic as well as a built-in function in R. HINT: use **1-sample Wilcoxon test** for the **difference** of income).

2. Use images `fi.jpg` (of the Faculty of Informatics) and `dms.jpg` (of Department of Mathematics and Statistics). Are the levels of green color the same over those two images?

   (a) Perform an exploratory data analysis of the problem. Create a histograms together with a kernel density estimations, boxplots and plots of the empirical distribution functions.

   (b) Formulate a hypothesis that could answer the posted question. Choose correct test statistics and test your hypothesis. Try different alternatives. HINT: use **Kolmogorov-Smirnov test** for comparing the empirical distribution functions.