

Name:

Master:

## Practice 2. Handling missing values

In this practice, you will use the dataset [ORGANICS](#). You will need to handle missing values.

Launch Visual Analytics or start a new report.

1. Explore the missing values from the loaded dataset using **Explore and Visualize** reports.
2. Create a pipeline using **Model Builder** to explore the missing values.
  - Indicate the missing values ratio of each variable in the dataset;
  - Explain the input variable data correlation;
  - Conduct a complete case analysis. Identify the complete cases.
3. Add to your pipeline a Variable Imputation block.
  - Explore all imputation method both for numerical values and categorical values;
  - Create single indicators for the categorical variables;
  - Create unique identifiers for the categorical variables and use them as input in the modeling process.
4. Add a **Logistic Regression** block to your pipeline.
  - Explore the binary target link functions; Create multiple **Logistic Regression** models with different binary target link functions and compare them.
5. Add a Neural Network block without any configurations.
  - Compare the models.

### Important:

- Answer the question is this practice either in 1-2 paragraphs or by adding a screenshot that proves that you have accomplished the assignment correct;
- Convert the obtained file to a PDF file and upload it to Moodle.

**Assignment due date: 01.05.2025 on Moodle.**