Plan:

1. Descriptive analysis + data visualization (exploratory data analysis)
   1. Data description:
      * ~~Problem description and research questions formulation: M~~
        + ~~What questions do we want to answer analysing the data~~
        + ~~What potential benefits may result from the analysis? (For example, the benefit could be: a better diagnostic method, better efficiency in detecting bad/good customers applying for a loan, separating groups of customers who can be targeted with a specific offer, identifying relevant features/variables, etc.)~~
      * Data characteristics: I
        + ~~Data size~~
        + ~~number of cases and features,~~
        + ~~types of features~~
        + information about missing values
        + information on unusual values
          1. non-standard coding of missing values
          2. non-standard binary values
   2. ~~Cleaning data:~~
      * ~~Data type~~
      * ~~Missing values as NA. How many?~~
      * ~~Add missing values (several methods and choose the best)~~
      * ~~Description M~~
   3. ~~EDA:~~
      * ~~Summary (min, max, quantiles, var) I~~
      * ~~Barplots I~~
        + ~~Balance~~
      * ~~Histograms M~~
        + ~~Distribution~~
      * ~~QQ Plot M~~
      * ~~Boxplots for int, cont M~~
        + ~~Outliers~~
        + ~~Initial assessment of discriminative ability of consecutive features (i.e. ability to separate objects from different classes).~~
      * ~~Barplots for categorises I~~
        + ~~Initial assessment of discriminative ability of consecutive features (i.e. ability to separate objects from different classes).~~
      * ~~Histograms for categorises M~~
        + ~~Distribution~~
      * ~~Correlations I~~
2. Classification along with detailed accuracy assessment
   1. Methods:
      * Linear regression I
      * K-NN I
      * LDA I
      * QDA I
      * ~~LR M~~
      * ~~Random tree M~~
      * ~~Bagging M~~
      * ~~Boosting M~~
      * ~~Random forest M~~
   2. Accuracy:
      * ~~Confusion matrix M~~
      * ~~Cross-Validation (CV) M~~
      * ~~Bootstrap-based methods I~~
      * ROC-curve I