

For the chosen dataset (multivariate, mixed attributes) from:

<https://archive.ics.uci.edu/ml/datasets.php>

perform:

1. **Check the quality of data** for further processing and **clean** if necessary (f.ex. null values, outliers....)
2. Using different **summary statistics** (adequate for continuous and categorical data) and **graphical tools** (from ggplot2, lattice or graphics packages) **in R language** such as:
 - histograms,
 - bar plots,
 - density plots, box and whisker plots,
 - stem-and-leaf,
 - missing plot,
 - cumulative distribution function plot,
 - QQ plots

for selected attributes (univariate case)

and

- correlation plots,
- scatterplot matrix, scatterplot matrix by class,
- density by class,
- box and whisker plots by class

for multivariate plots

**try to FIRST understand your data,
then**

**describe the descriptive analysis results for the selected most important
attributes (5 pages)**

and formulate assumptions and hypothesis for further modeling !!!!