

PD Estimation:

Link:

<https://github.com/INTQuant-Katowice/2024/tree/main/Data>

Input data:

- DataPD.txt
- Description.txt

Tasks:

- (1) Load datasets
 - Verify dataset was properly loaded
 - Display excerpts of the datasets and check if loaded properly
- (2) Analyze datasets:
 - Count number of NAs
 - Compute quantiles, mean, std deviation
- (3) Plot distributions
 - Plot histogram of a selected variable
 - Plot histograms of all variables (adjust for outliers)
 - Compute correlations and plot correlation matrix
- (4) Perform regression and estimate the model
 - Clean data
 - Perform regression
 - Check basic outputs
 - Prepare ROC curve and calculate AUC
 - Calculate probability of default for three given institutions:
 - Bank ABC (ID = 484)
 - Bank XYZ (ID = 47)
 - Bank QQQ (ID = 2741)
- (5) Improve on the regression...