

Part A - Preprocessing

You are given a data set that contains details of a bank's customers. The task here is to use the preprocessing techniques that we've shown in the class to clean and transform this data set.

1. Are there duplicates? If so, remove them.
2. Uni-variate analysis:
 - a. Type of each feature
 - b. Mean, median, quantiles for continues
 - c. Histogram or frequency tables for categorical
3. Outliers:
 - a. Identify the outliers
 - b. Choose the correct treatment
4. Missing values:
 - a. Identify missing values
 - b. Impute missing values with your preferred way
5. Transform categorical variables with your preferred way
6. Scale the data set

Note: save a version of your data after step 4 for Part B.

Deadline: November 15th