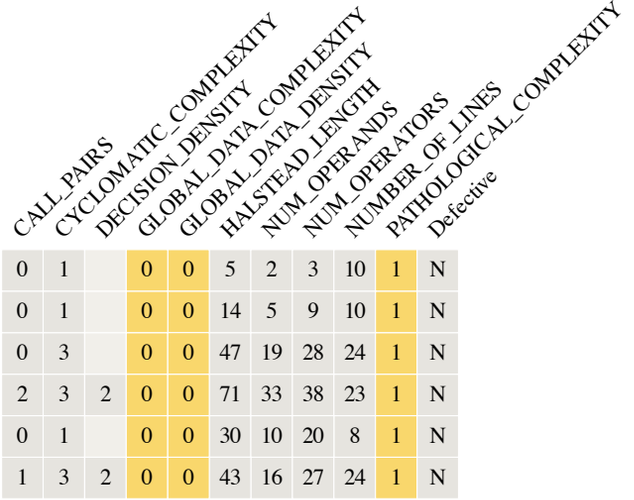
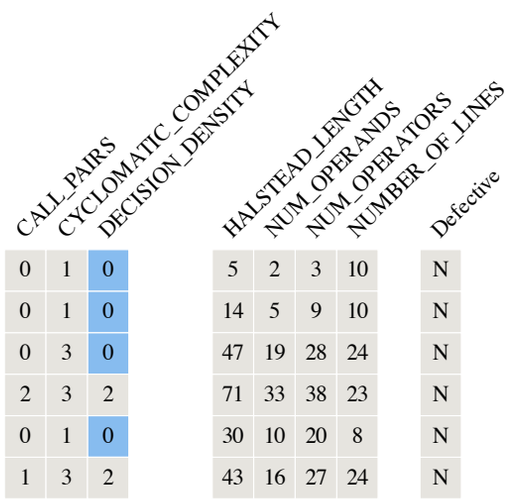
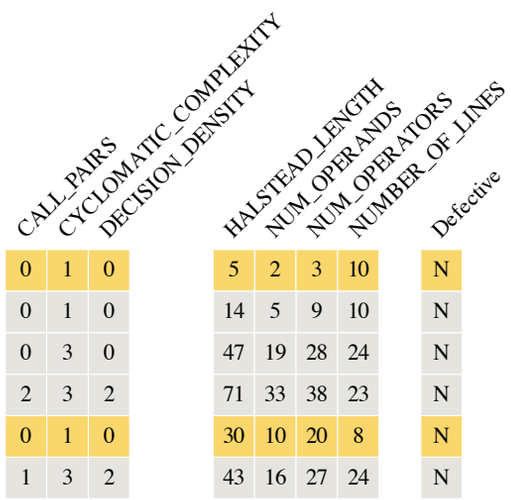


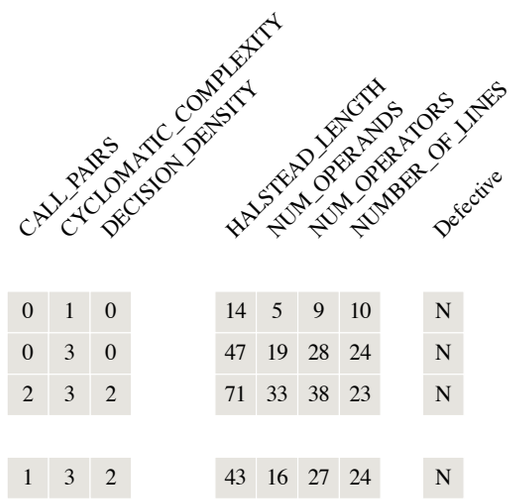
Remove columns that have the same value for every row because they do not provide any information for modelling.



Replace missing *DECISION_DENSITY* values with zero. Based on other MDP datasets without missing *DECISION_DENSITY* values, one can deduce that they likely occurred due to a division by zero error and can be replaced with zeros.



Remove rows that are duplicates of other rows to assure models are tested on unseen data only. Also remove rows that are inconsistent, meaning all column values are the same except for the class label (one is classified as defective and the other is not).



The CM1 dataset is now ready for use in modelling.

Deleted Edited Unchanged*

*A lighter value indicates a missing data value