

Practice2A (Wine Datasets)

Dataset

Data used: red + white CSVs wine merged

Preprocessing

Categorical variables been cast to bool or made ordinal

Outliers

Dropped with IQR

Missing Values

Were held with Scaling + kNN imputer

Visualizations

Different visualizations (boxplots, histplots, scatterplots, bar plot and heatmap)

Result

Ready to proceed train + test datasets with clean data