* Check missing values by looping through each column

[col for col in data.columns if data[col].isnull().sum() >0 ]

* Split train and test set and examine only train test
* Remove constant features from dataset (using Variance Threshold)
* Remove quasi-constant features (99% same values)
* Drop duplicated features
* For correlation select numerical features only and remove correlated features
* Remove features with univariate roc-auc <=0.5
* Use wrapper methods forward/backward or exhaustive

**USE EITHER OF THEM**

* Use Lasso to get important features
* Tree Importance
* If using linear models then use importance generated by regression coefficient or Lasso
* RFE with tree, takes lots of time not better than SelectFromModel