* First Step is to remove the features which has more than 60 percent missing values
* Encode the categorical features
* Splitting the data (<https://stackoverflow.com/questions/56941740/in-outliers-detection-train-test-split-after-or-before-fit-the-algorithm>)
* Second step is to check for outliers and remove them  
  (<https://datascience.stackexchange.com/questions/39751/which-order-is-correct-feature-selection-then-outlier-detection-or-vice-versa#:~:text=1%20Answer&text=In%20majority%20of%20the%20cases,be%20done%20after%20outlier%20detection.&text=Many%20feature%20selection%20algorithms%20like,to%20remove%20the%20outliers%20beforehand>.)
* Impute Missing Values