We use BCor-SIS, a generic non-parametric sure independence screening procedure based on ball correlation, is able to pick out explanatory variables related to response. The linear, non-linear or linear interaction effect relationship can be captured by BCor-SIS even though data is heavy tail or existing outliers. More importantly, BCor-SIS is able to retain all of the important features in the model with probability tending to 1 under mild conditions. Ball correlation is defined as follow in [reference]

The detail procedure has been programmed in the R package Ball, so we utilize the core function \*\*bcorsis\*\* to find out several X which are correlated to CA's Y2. Then we ues the stepwise procedure to narrow down the range of X. Finally we establish a simple linear regression model indicating the relationship between Y2 of CA and GOCCB.