

# DONATIONS

A SIMULATION TO LEARN PREDICTIVE ANALYTICS



1



## EXPLORE AND ANALYZE THE DATA

Range, variable types, correlation, null values

### LAST YEAR INFORMATION (GAVE, AMOUNT)

History excluded Last Year

Duplicate Data to train the model

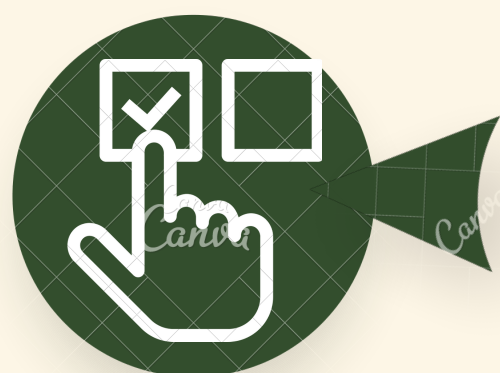
Last Year

History contained this data

This Year

History needs to be updated with last year information

## PREPROCESSING



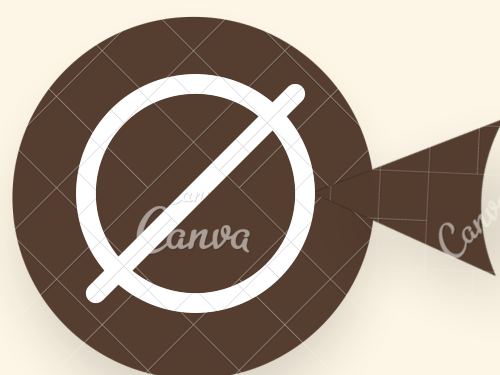
### Selection of Variables & Dummies

Remove:

- Name and LastName

Dummies:

- Education, Cities



### Fill NULL values

Recency, Frequency, Seniority, TotalGift, MinGift, MaxGift



### Duplicate data

First Part: Last Year Data

Second Part: This Year Data

Update variables if they gave last year (Frequency, Recency)

## PREDICT AMOUNT

Predict amount that each person would donate

### Best Model:

XGBoost  
MAE: 61.73



### Surplus

\$11,705,850.00

### Expenses

\$1,800,000.00

#### Feature importance

Salary			193
TotalGift		114	
Age		102	
Woman	46		

## PREDICT PROBABILITY

### New Variable: Contact

Predict probability that a person would donate with contact and without contact

Best Model:  
Decision Tree  
Accuracy: 86



#### Feature importance

Contact			136
NbActivities		108	
Recency		90	
Referrals		84	

## UPLIFT

ProbContact -  
ProbNoContact \* Amount



Order By:  
ProbContact  
Uplift  
Prediction

Operating surplus

\$12,315,510.00

Expenses

\$2,536,425.00