Hello!

I am Rupali Roy

I a final year graduate student at UT pursuing masters in Information Science with a focus on Data Science and Machine Learning! I am here to give my final year's capstone project presentation.

Exploring Data Segmentation and Donation Giving Behaviors

By - Rupali Roy

As the old proverb goes:

“If you want happiness for an hour, take a nap.
“If you want happiness for a day, go fishing.
“If you want happiness for a year, inherit a fortune.

“If you want happiness for lifetime, help somebody.”

What motivates people to donate?

1. People donate to feel better

2. They believe they can make a difference

3. People give to belong

4. Tax Deductions

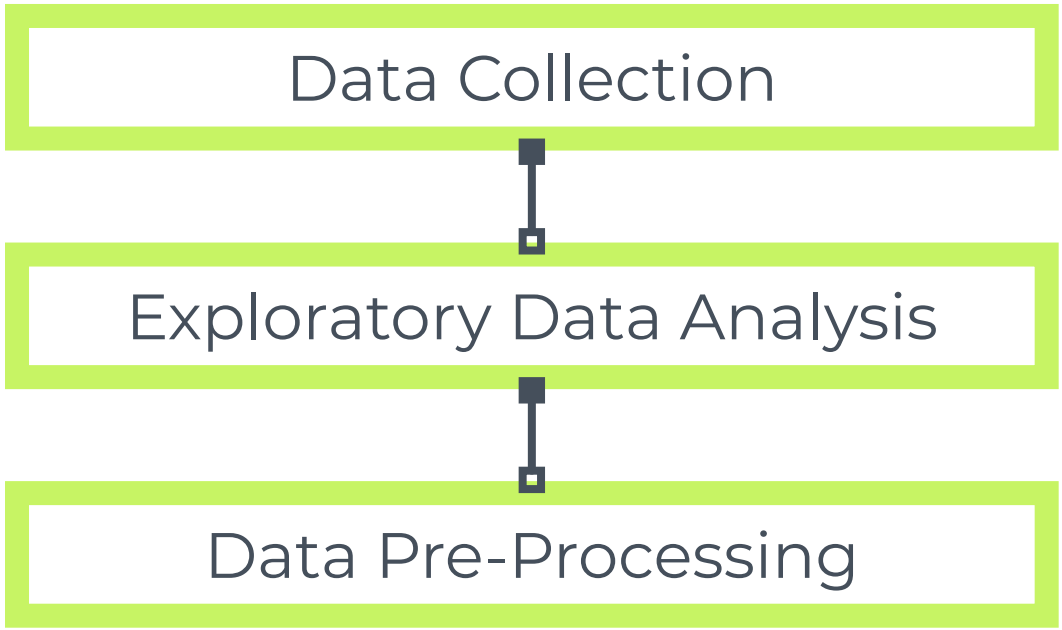
1.

Problem Statement

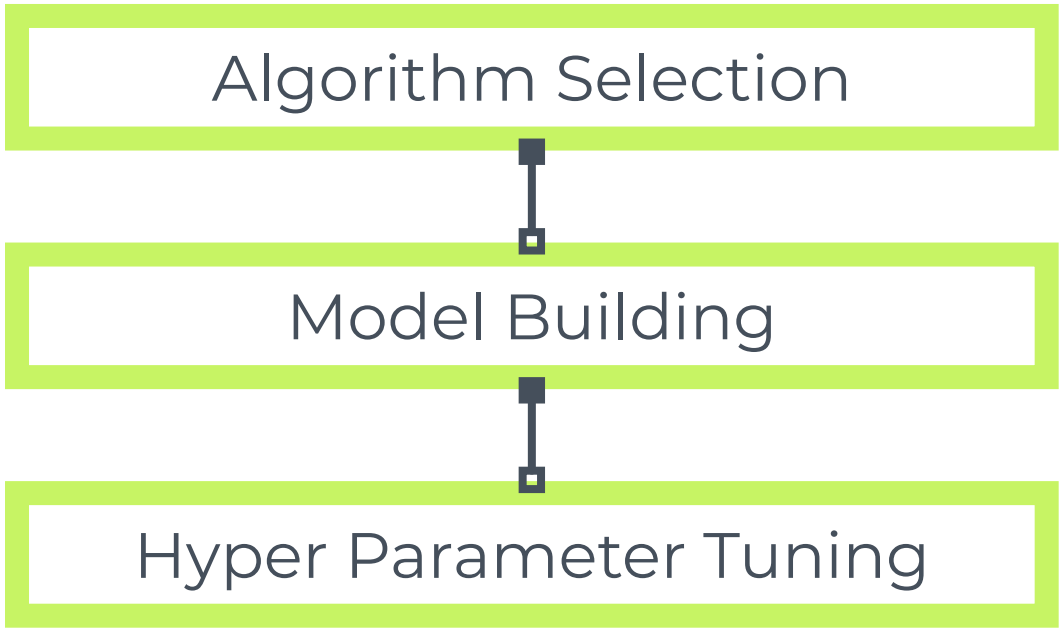
1. Understand Donation Giving Behaviors

2. Build a predictive model for future donations

Data Exploration



Model Building



Data Collection

- Bio_Base_Data - Demographics Data
- Gift_Total_by_CSU_Since_2015
- Dictionary

Exploratory Data Analysis

1. Check for meaning of the fields ced by the combination
2. Check for the data structure
3. Check for duplicates
4. Check for missing values
5. Check for field formats
6. Check for zero values that shouldn't be there e.g. age, text, year
7. Decide which fields are nominal (categorical), ordinal, dummy, interval or ratio
8. Check for levels of categorical variables
9. Check for outliers
10. Check for inconsistencies
11. Check for hyphen , commas, or period in the wrong places

Handling Missing Data

AGE 20495 - 22%

PROFESSIONAL_SUFFIX 83865 - 92%

GENDER 473 - 1%

GENERATION 20495 - 22%

HH_CAPACITY_RANGE 46782 - 51%

EX_MEMBER 319 - 0%

ETHNICITY 41541 - 45%

MARITAL_STATUS 2305 - 3%

PREF_STATE 1434 - 2%

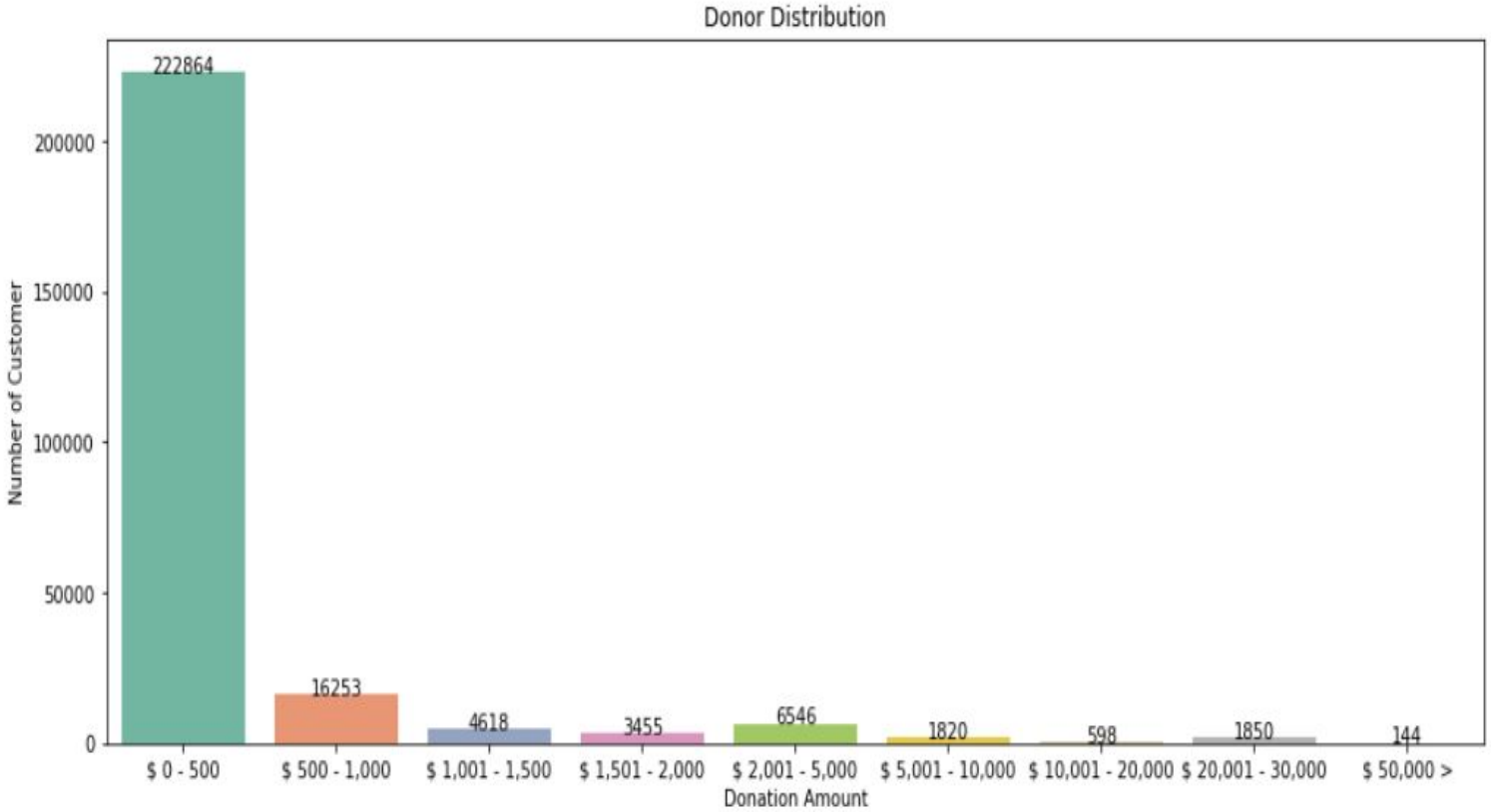
PREF_ZIPCODE 507 - 1%

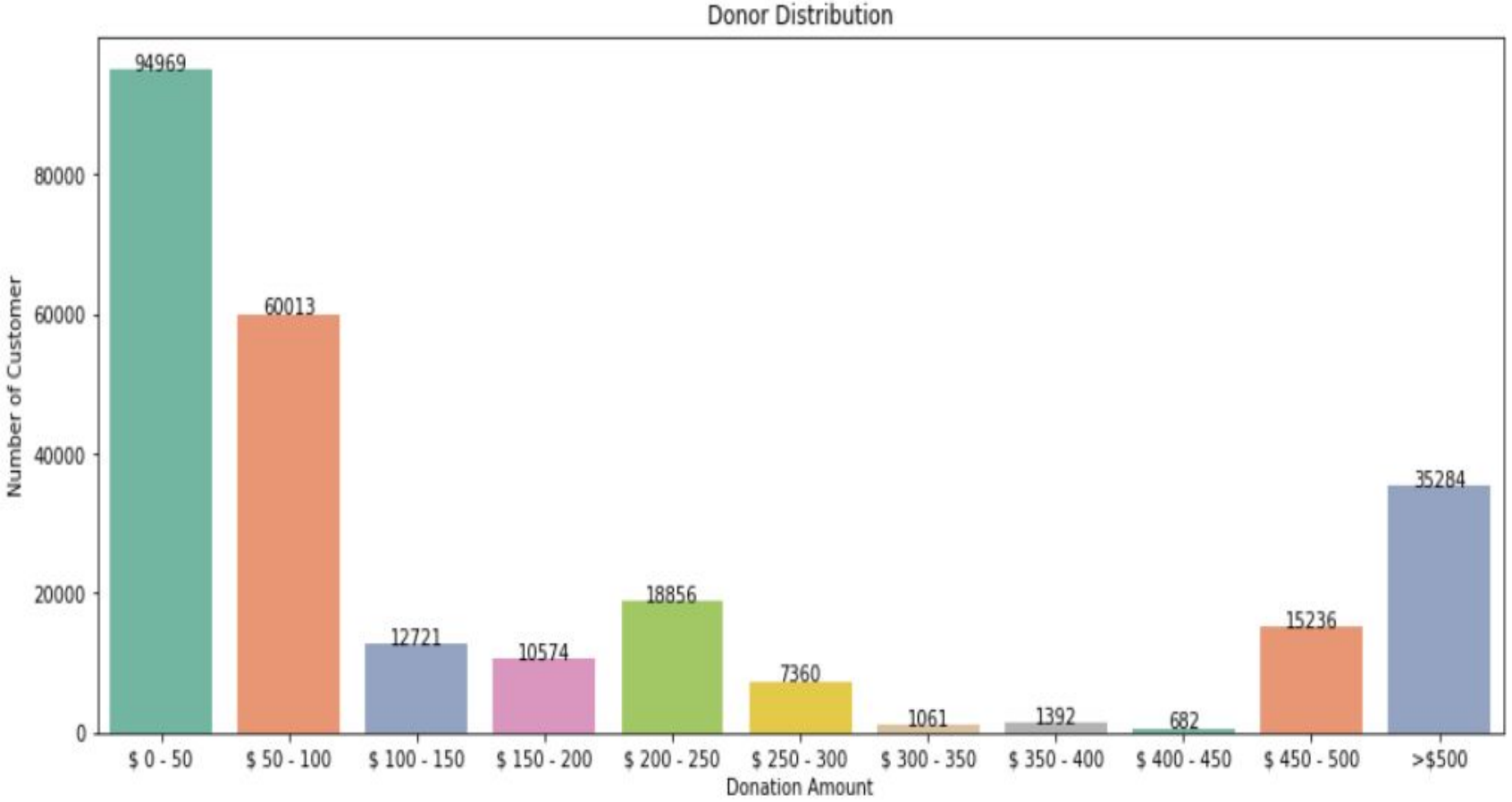
FIRST_DEGREE 33309 - 36%

SECOND_DEGREE 80592 - 88%
FINAL_DEGREE 90551 - 99%

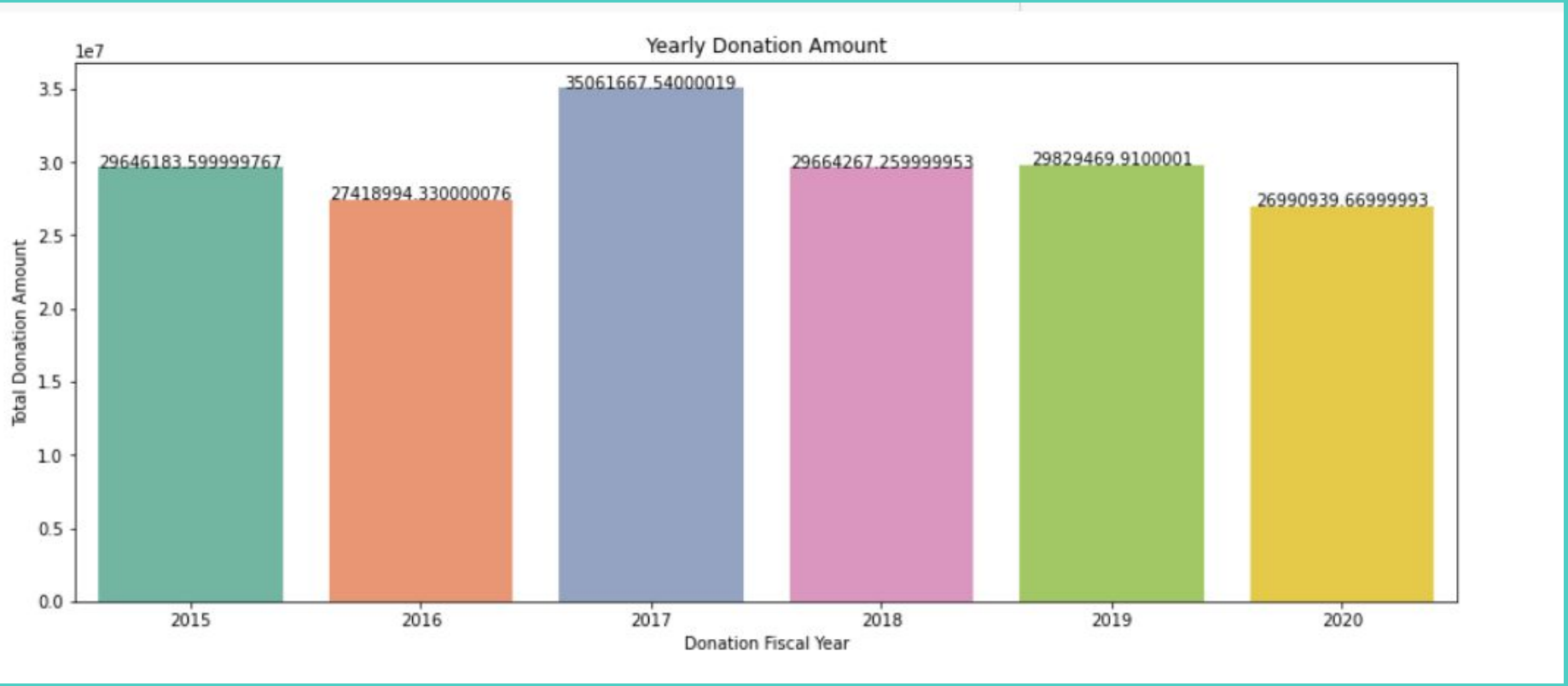
Imputation By:

1. Mean
2. Mode
3. K - Nearest Neighbors
4. Drop Na values

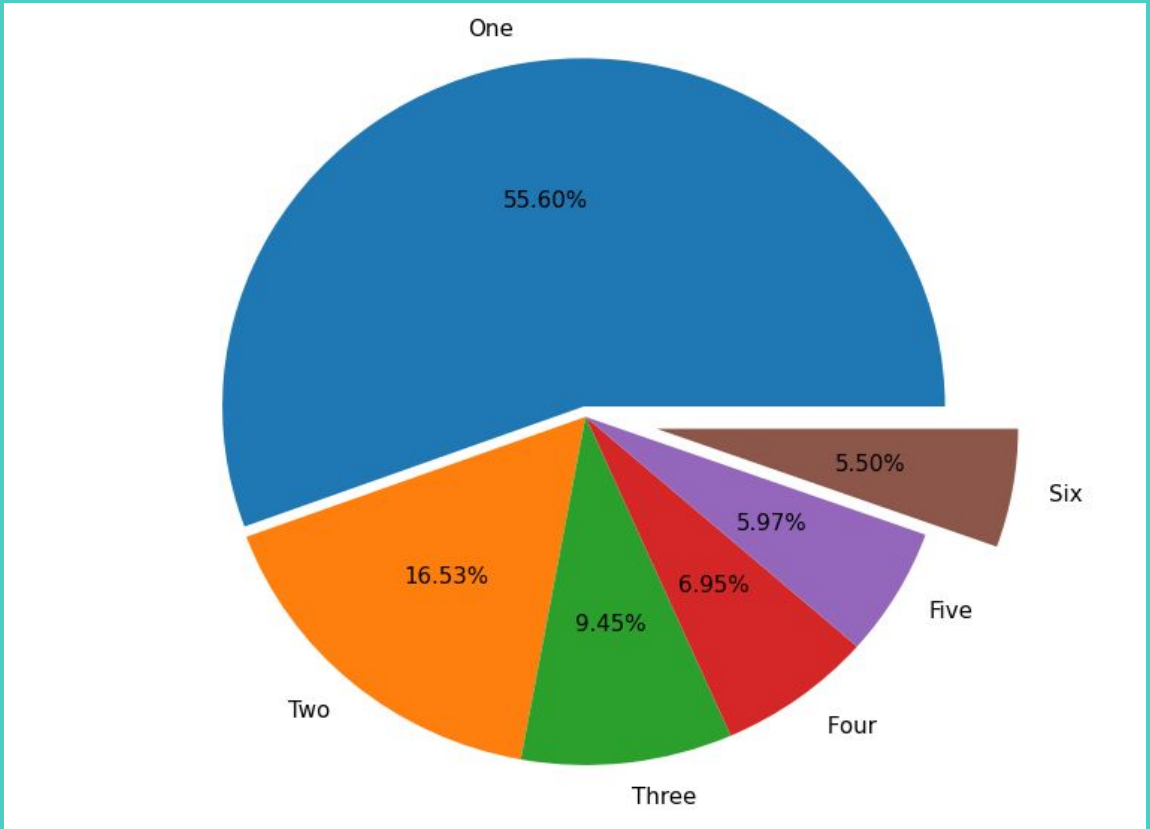

Donor Distribution
Number of Customer
0
50000
100000
150000
200000
222864
16253
4618
3455
6546
1820
598
1850
144
$ 0 - 500
$ 500 - 1,000
$ 1,001 - 1,500
$ 1,501 - 2,000
$ 2,001 - 5,000
$ 5,001 - 10,000
$ 10,001 - 20,000
$ 20,001 - 30,000
$ 50,000 >
Donation Amount


Donor Distribution
Number of Customer
0
20000
40000
60000
80000
94969
60013
12721
10574
18856
7360
1061
1392
682
15236
35284
$ 0 - 50
$ 50 - 100
$ 100 - 150
$ 150 - 200
$ 200 - 250
$ 250 - 300
$ 300 - 350
$ 350 - 400
$ 400 - 450
$ 450 - 500
>$500
Donation Amount

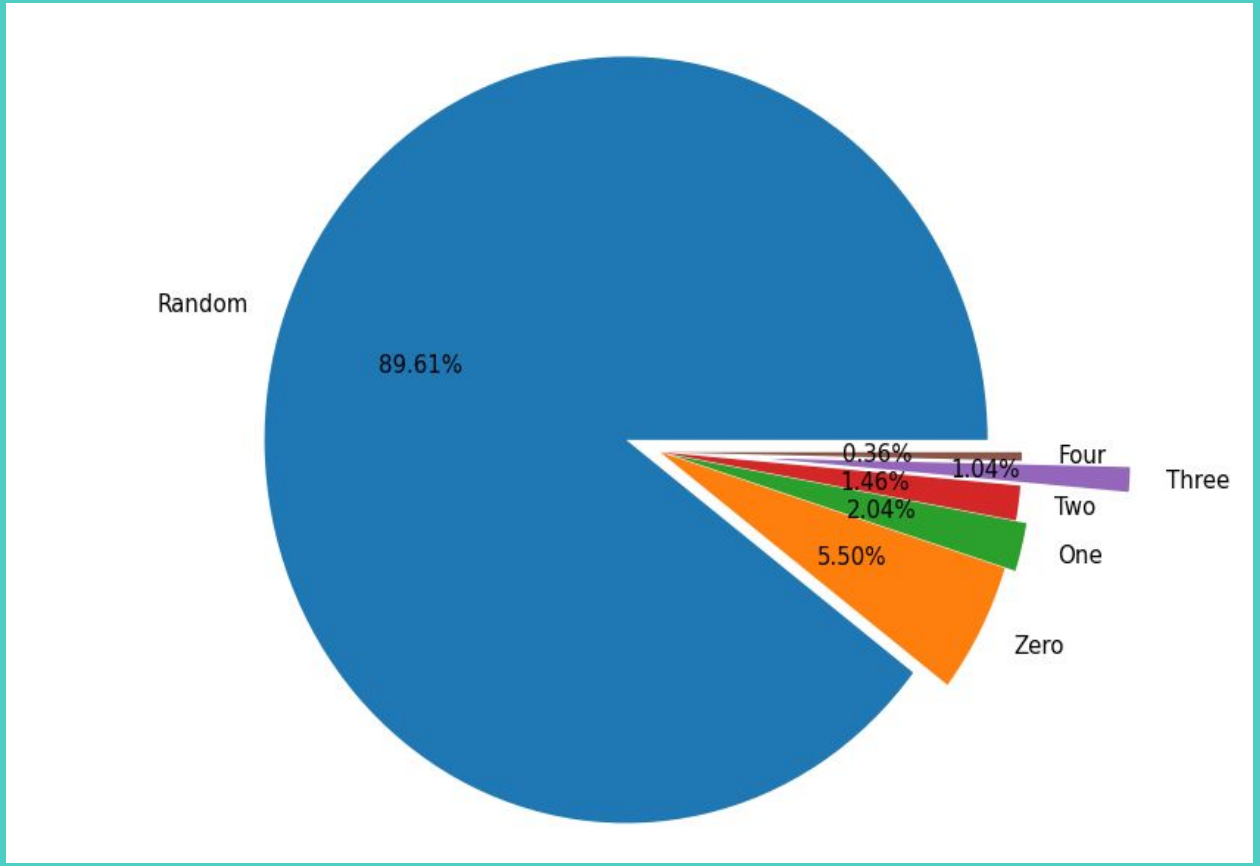
Total Donations declined in 2020



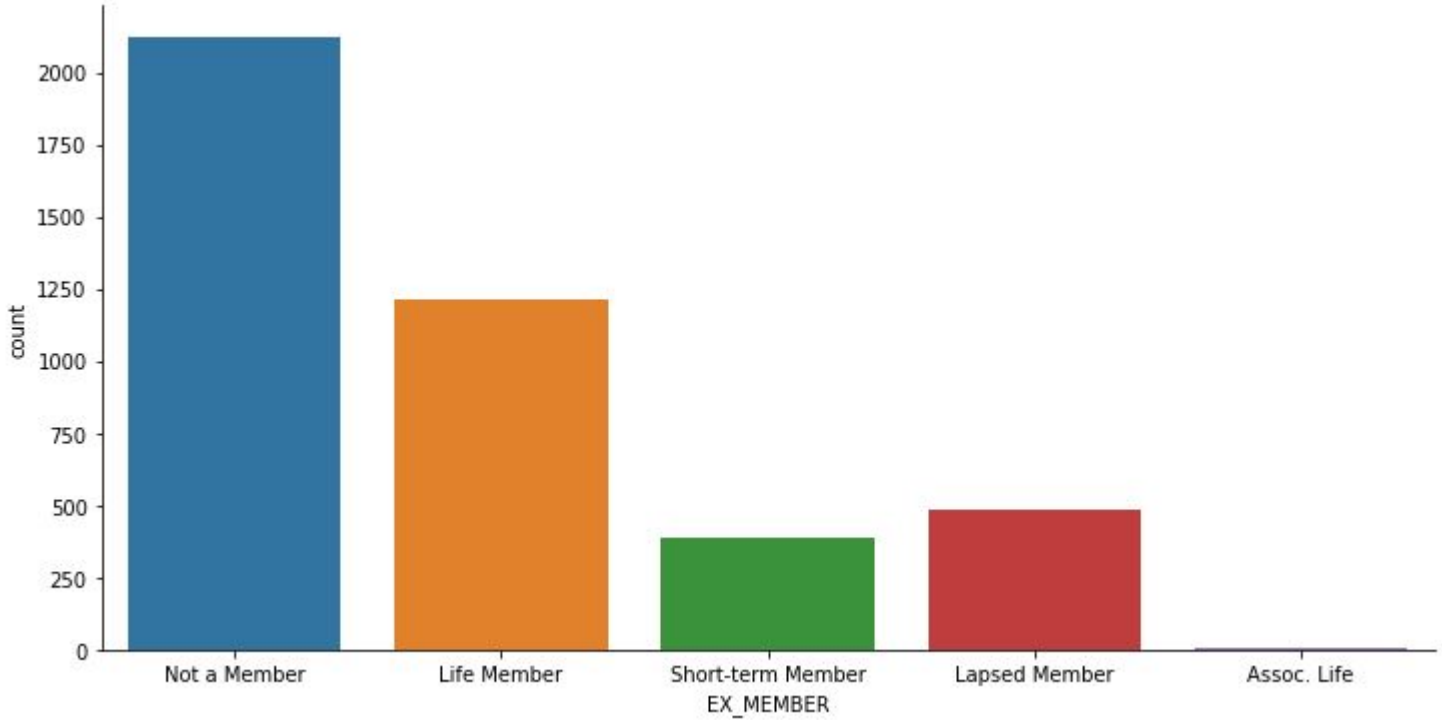
Only 5.5% of people donated every year



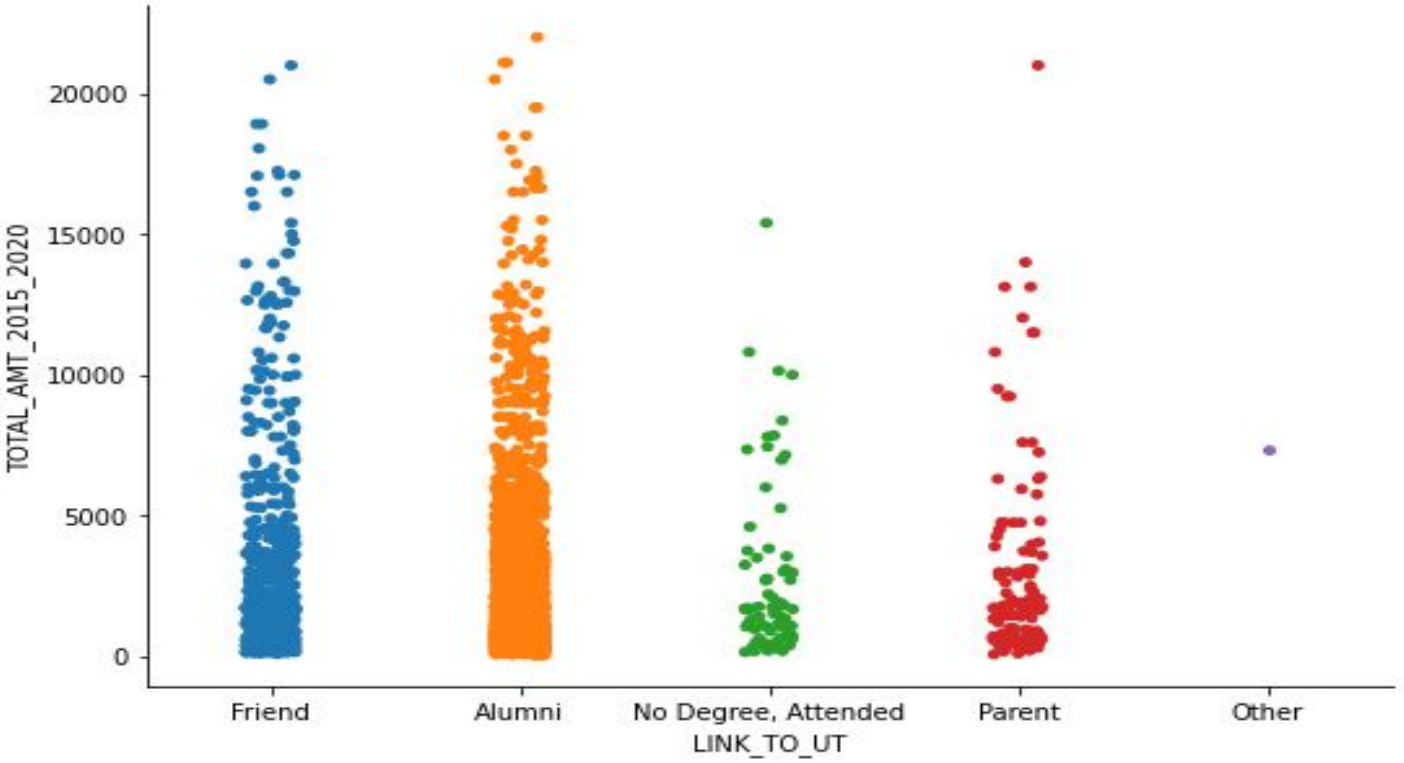
89.61% of people have random giving pattern



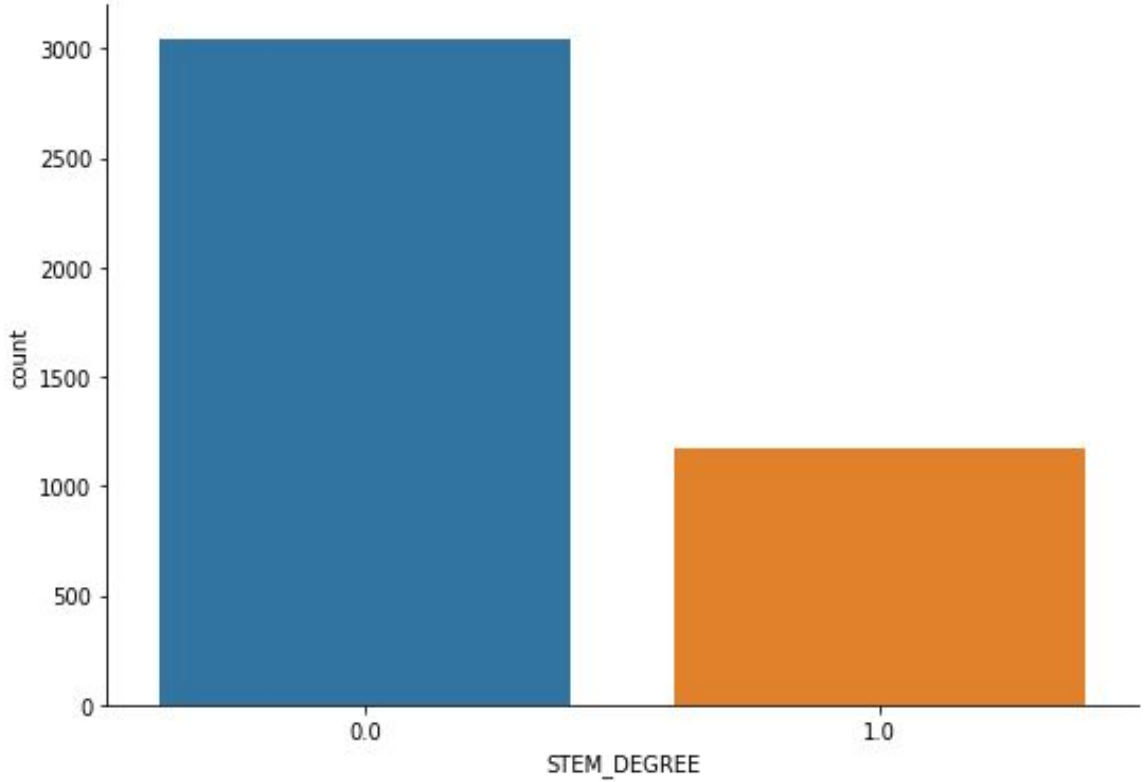
Giving Behavior with skip years=0



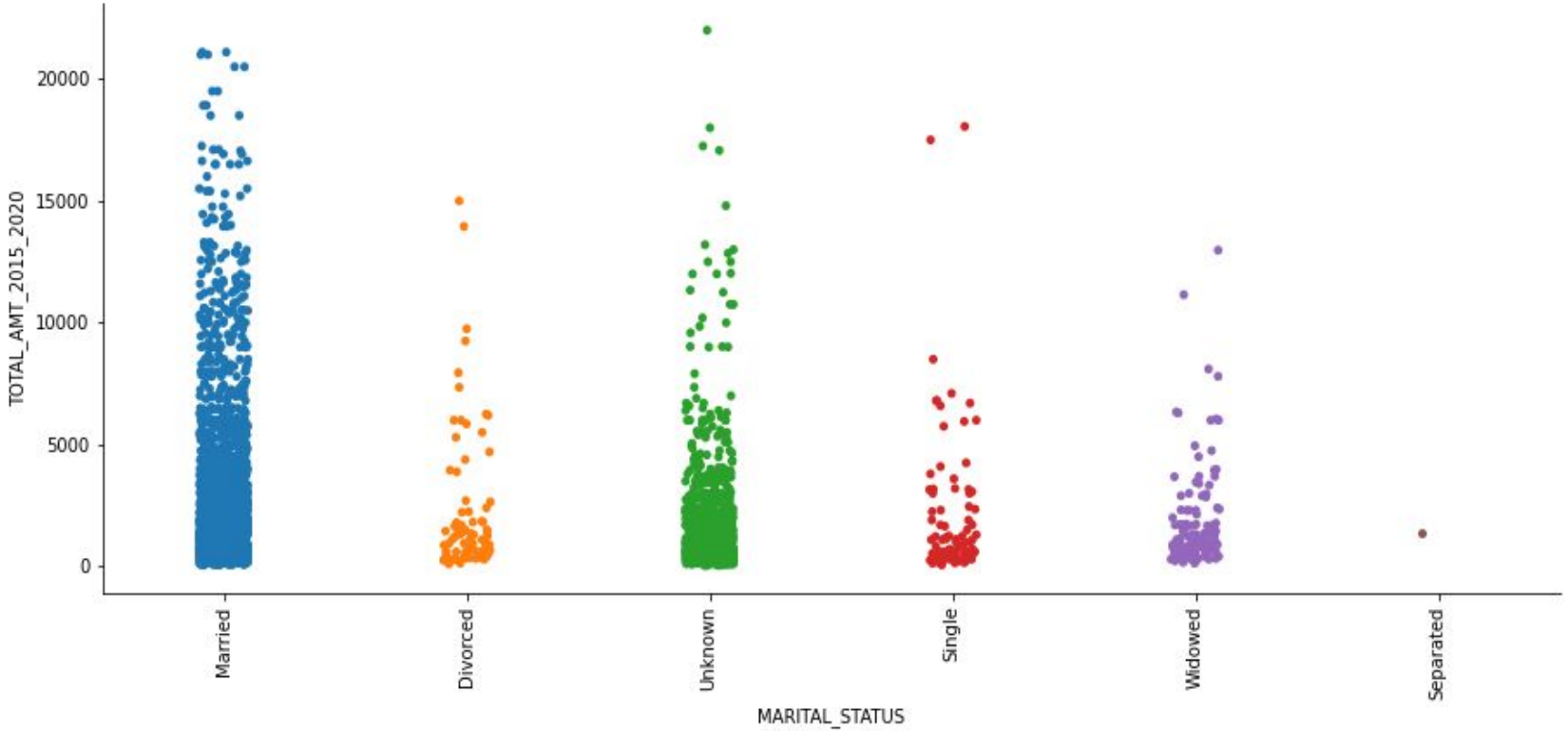
Most of the people who donated regularly are “Not a Member”



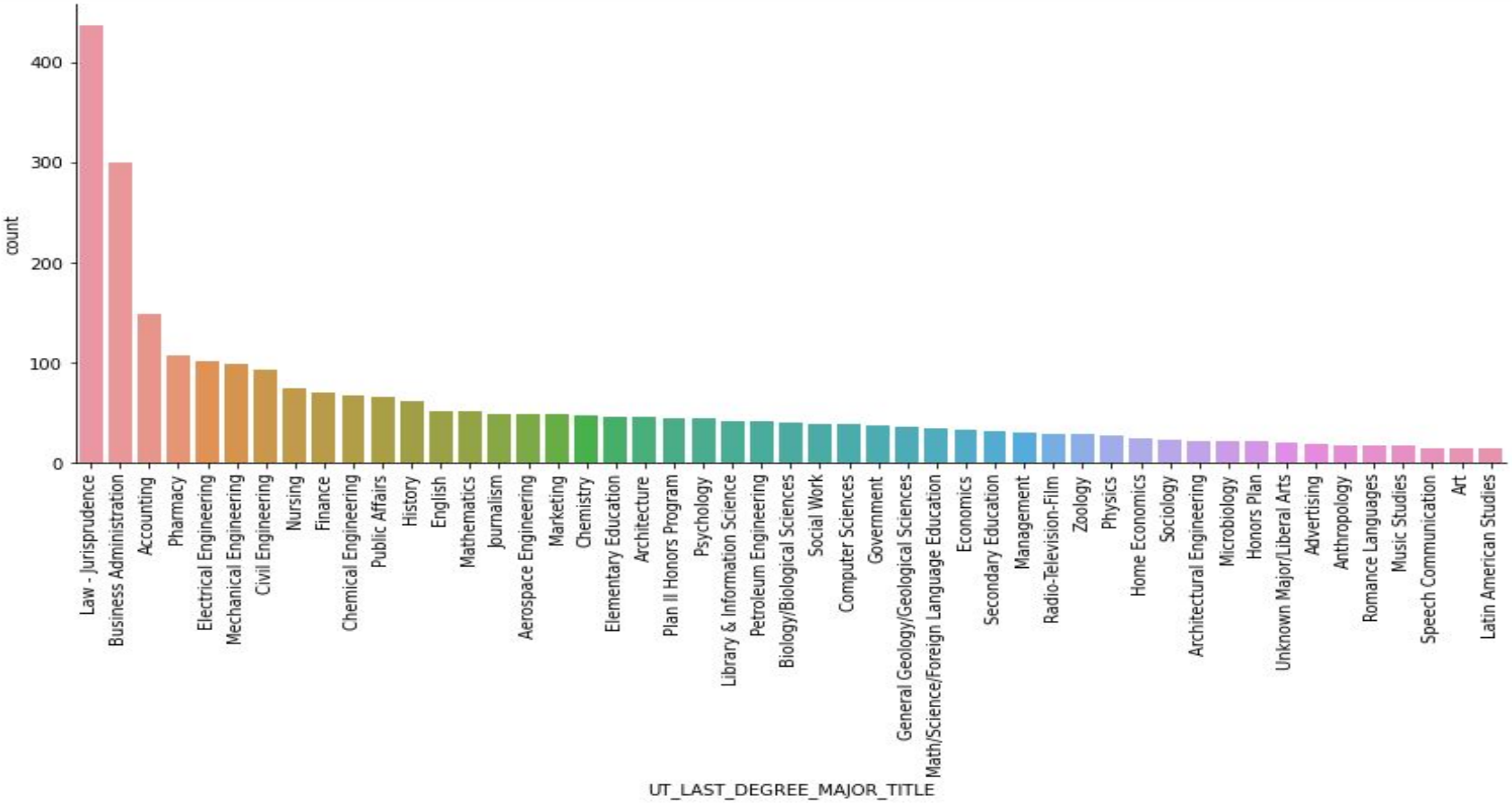
“Friend” and “Alumni” are most likely to donate



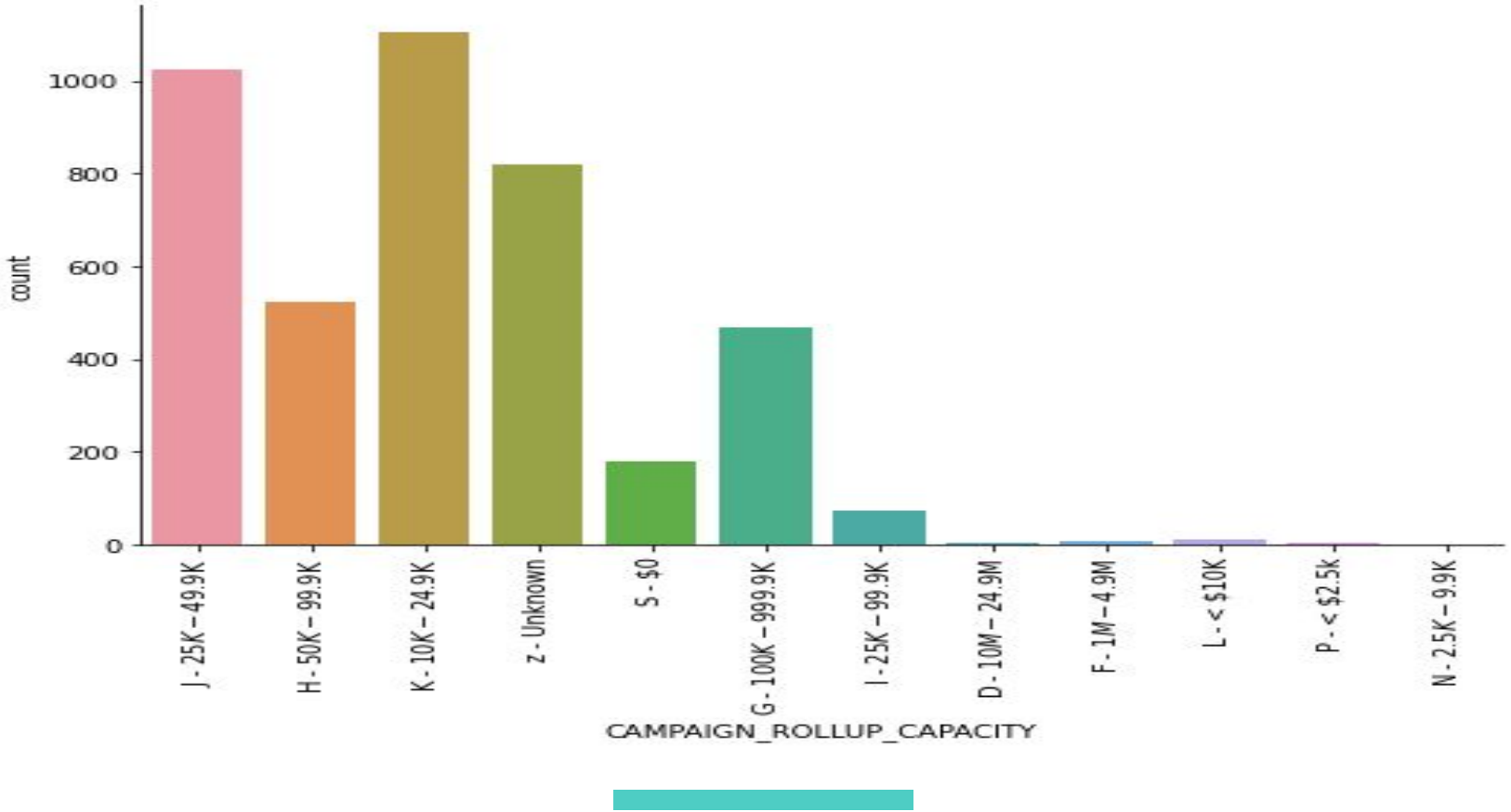
Most of the Donors don’t have a “STEM” Degree



People who are “Married” are more likely to donate!



“Law Jurisprudence” and “MBA” have major contributions every year!



Majority donors belong to K - 10K to 24.9K Campaign Rollup Capacity

Model Building

Approach

Binary Classification Problem to detect if a person will donate next year using the demographics data and the previous years donation giving behavior

Machine Learning Algorithm

Feature Variables

Total Input Features : 57

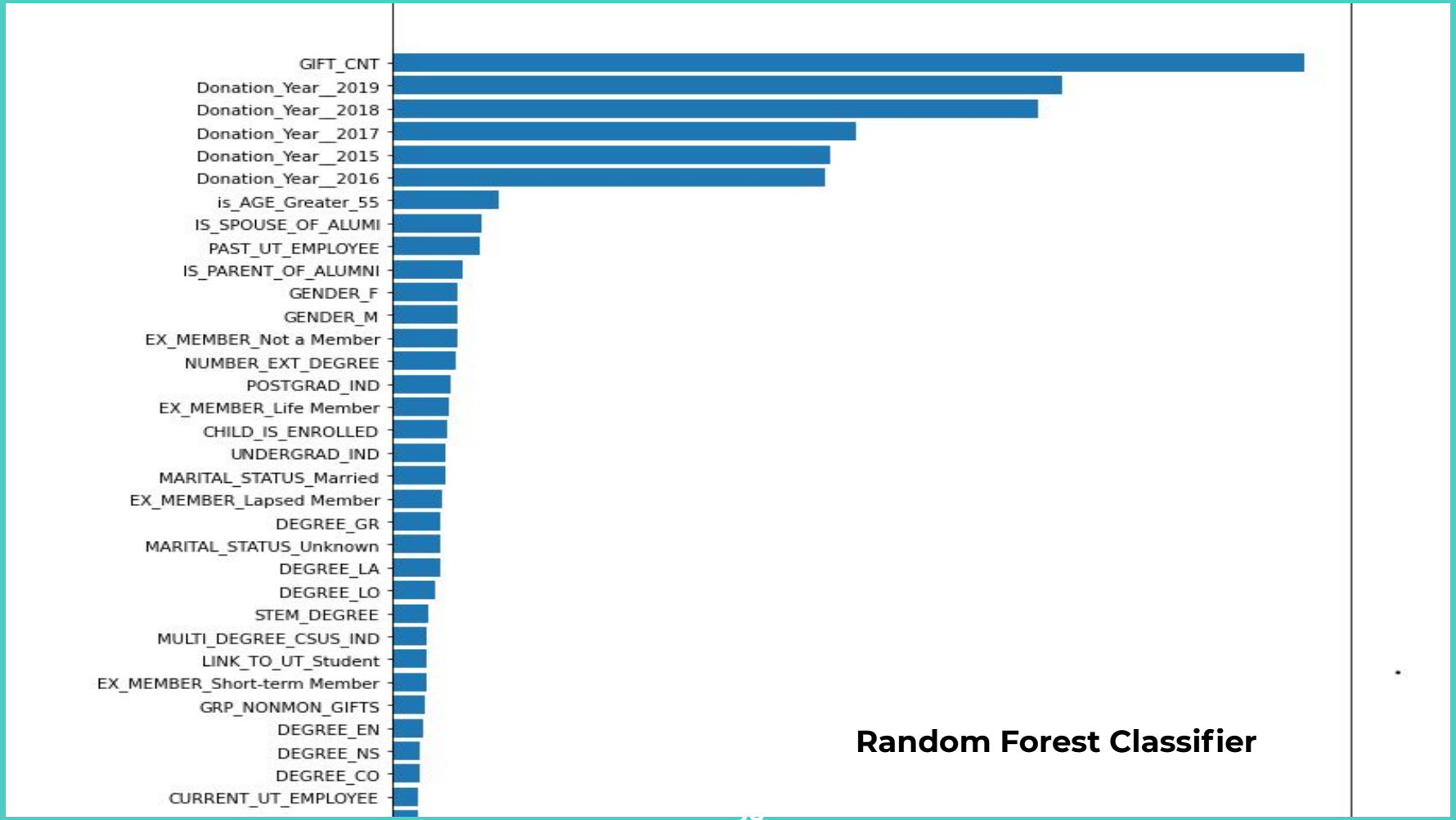
Output Variable : Donate - 0 or 1

Evaluation Metric used is ROC_AUC score

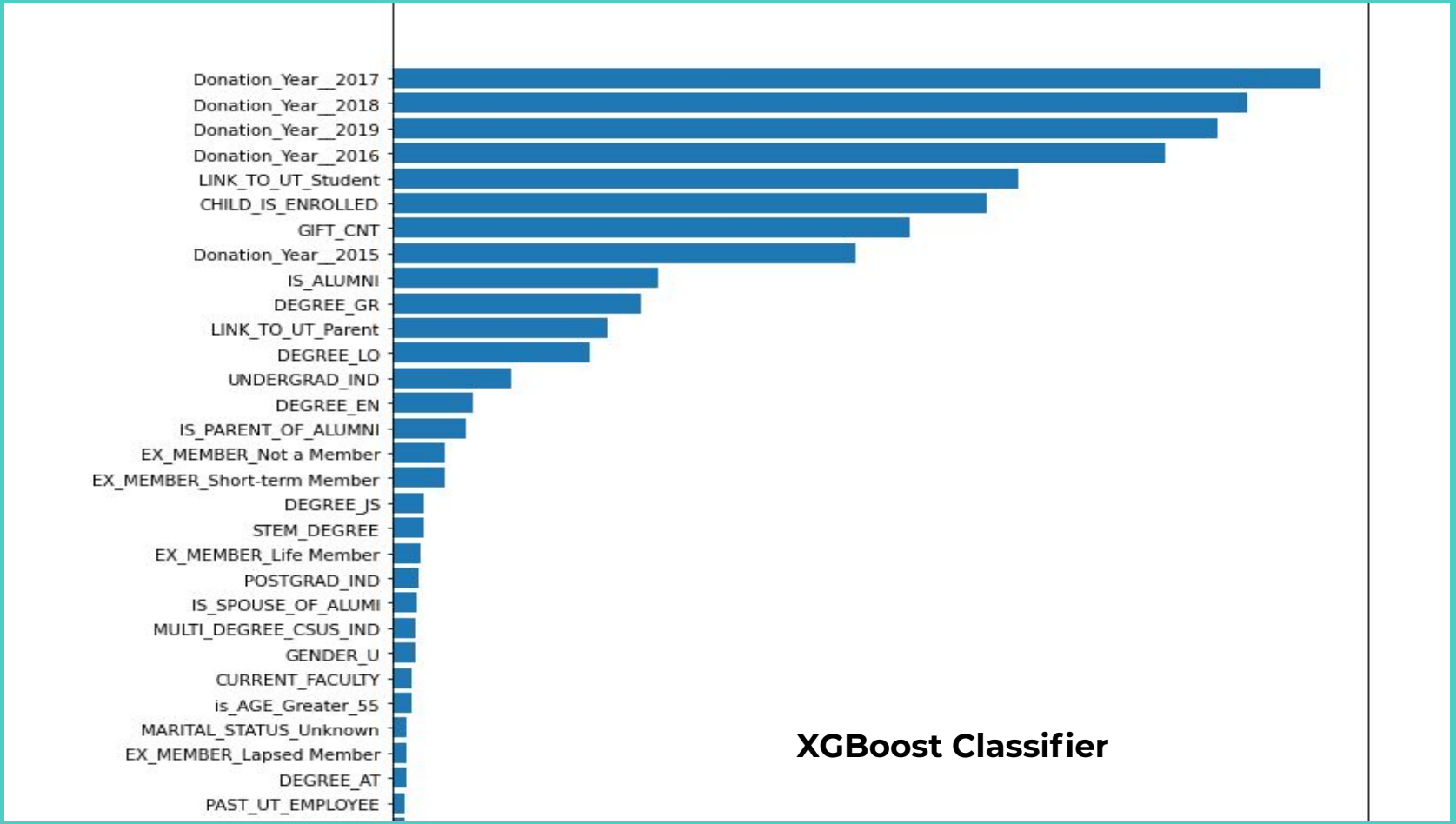
Evaluation Metric is Precision and Recall & ROC_AUC score

Model	Precision	Recall	F1 Score
Logistic	0.72	0.75	0.75
Decision Tree Classifier	0.83	0.83	0.83
Random Forest Classifier	0.85	0.86	0.86
XGB Classifier	0.87	0.87	0.87

XGB Classifier performed the best with ROC-AUC score of 0.92



Random Forest Classifier



XGBoost Classifier

Important Features!

1. Previous Years Donation Giving Behavior
2. Link_to_UT_Student
3. Child_isEnrolled
4. Gift_Cnt
5. Degree_EN
6. Is_Alumni
7. Link_to_UT_Parent

Ideas!
Campaign focused on the important features!

Thanks!

Any questions?

You can find me at
rupali.roy@utexas.edu

Special thanks to Ang Siew for helping and guiding me throughout the project!