Direct Social Communications

your partner in fundraising

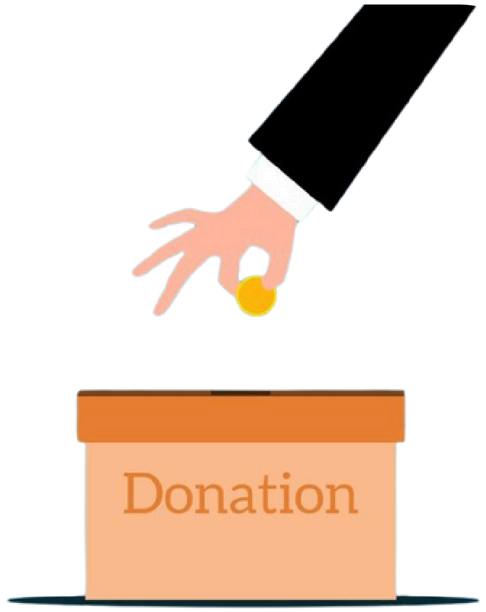
DSC: PREDICTING DONATIONS FOR FUTURE CAMPAIGNS

Know your Donors, Grow your Impact

ABOUDI Arajem | AZUBUIKE Olive |VARGAS PEREA Angie Tatiana

TABLE OF CONTENTS

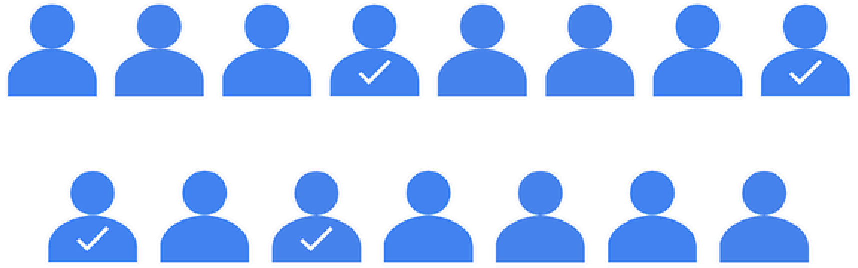
THE BUSINESS PROBLEM



DSC is Belgium's only specialized full-service **fundraising agency**, since 1985, by conducting detailed analyses of data and donor behavior, **ensures optimal fundraising results**.

BUT

latest campaign has had a **low performance** due to **RANDOM SELECTION**



GOAL

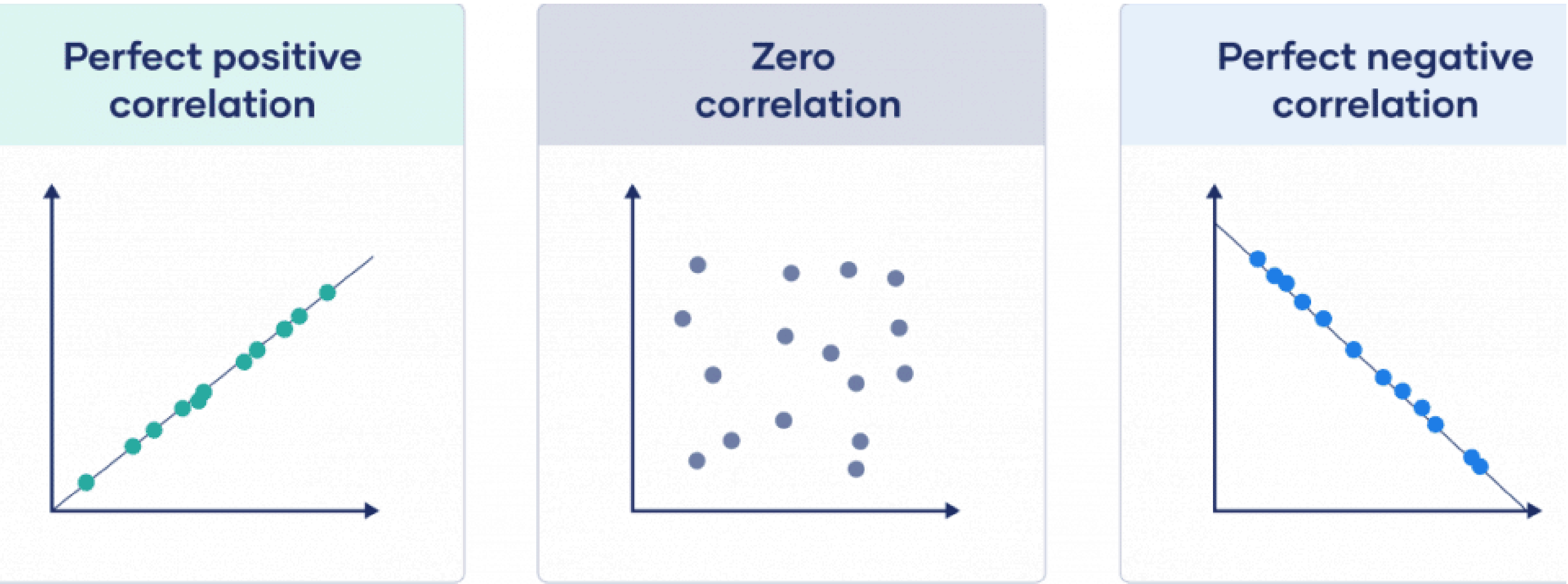
Identify donors who, if contacted, will donate more than average.

DATA PRE-PROCESSING

Process done for :
DONORS
GIFTS
CAMPAIGN

Cleaning Data

- MISSING VALUES
- NULL VALUES
- DATA CONVERSION
- OUTLIER TREATMENT



Source: Scribbr, *Correlation Coefficient*

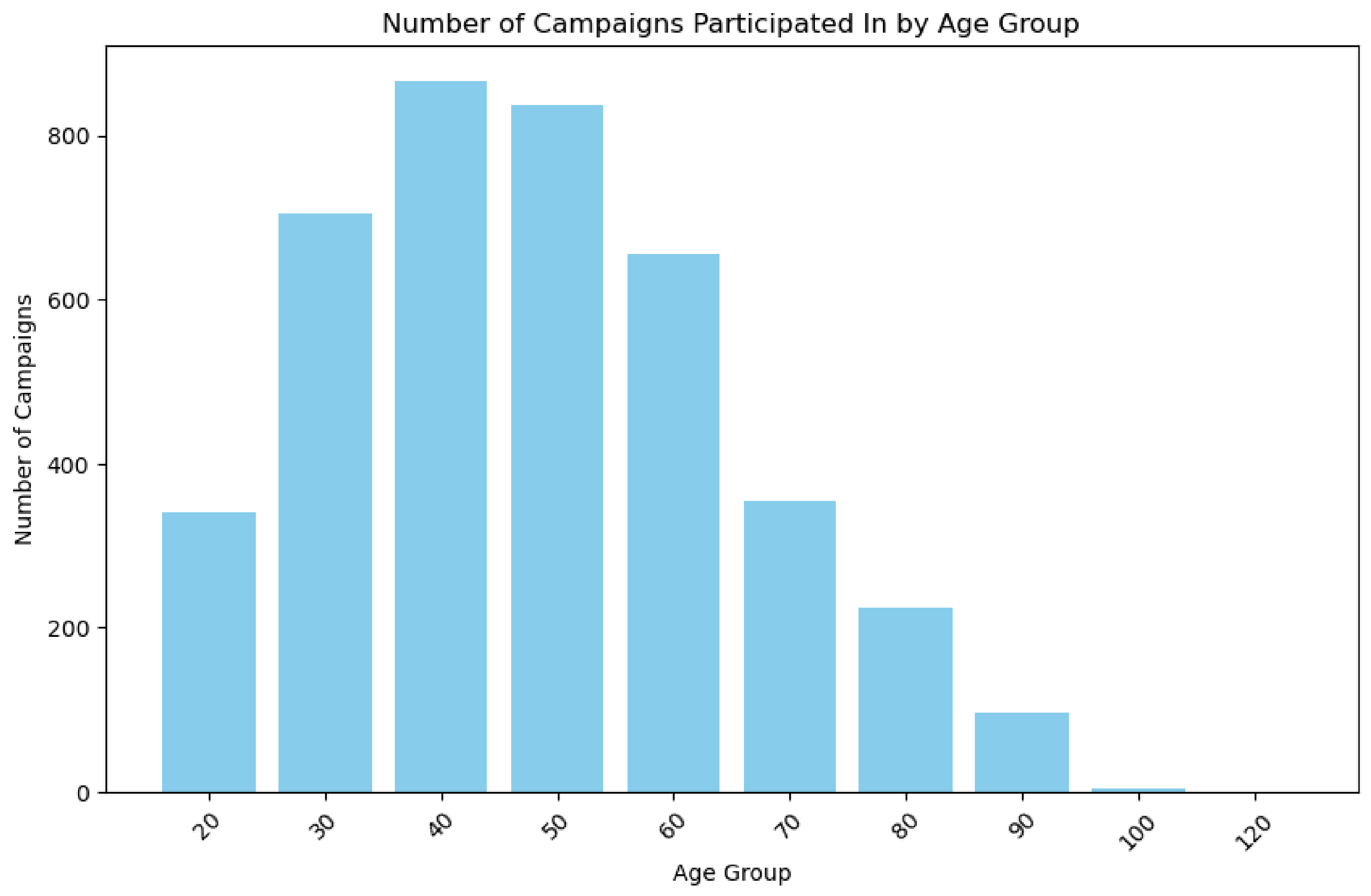
```
.is.null.sum()
.describe()
.info()
.dtypes
```

DATA & FEATURES CREATION

Historical Data

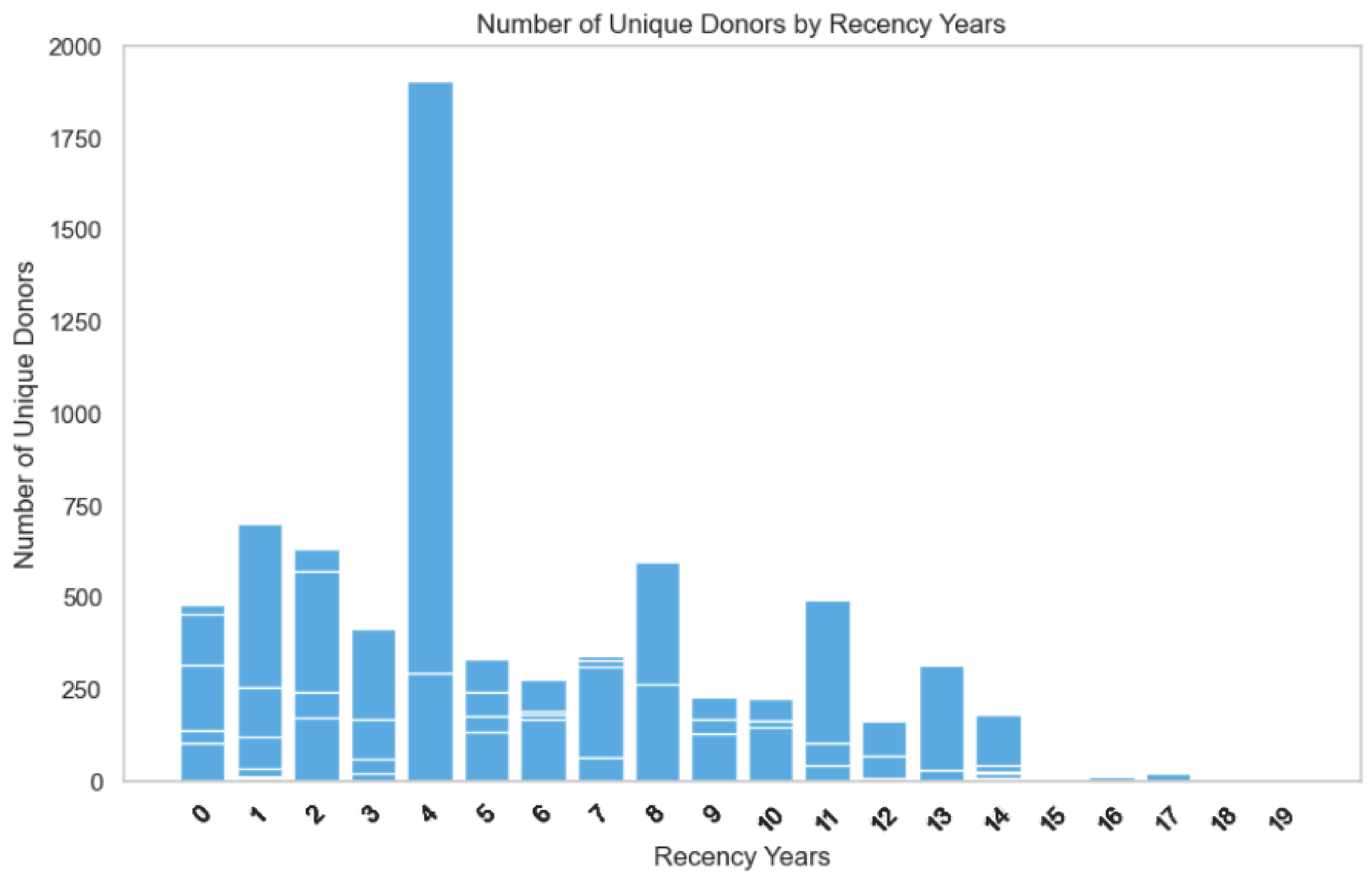
ALL THE PERIOD & LAST YEAR

- **Total amount** of donations
- **Average amount** per donation
- **Frequency** of donations
- **Max amount** of donations
- **Min amount** of donations



Historical Data

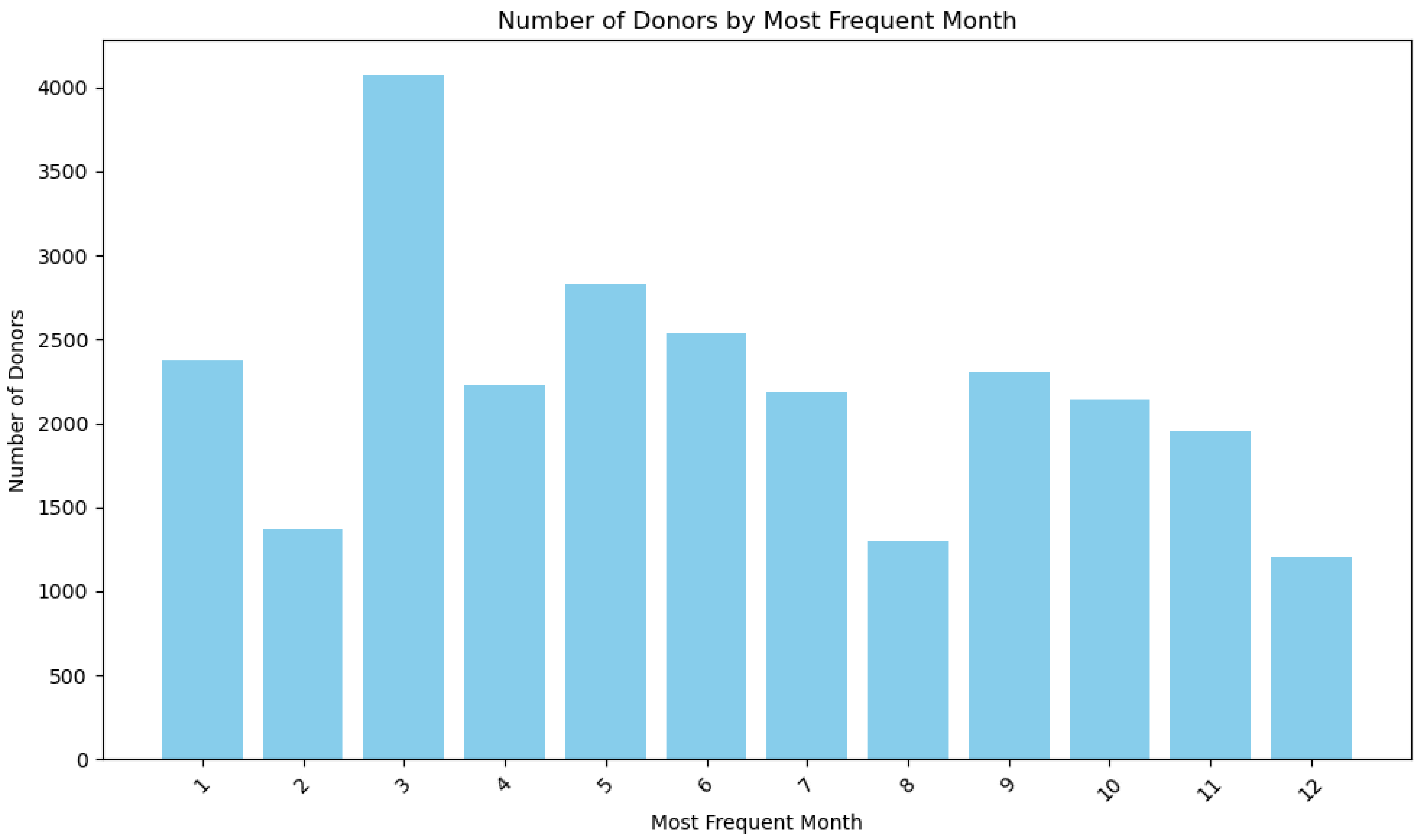
- Length of relationship
- **Recency**
- Donor lifetime
- Average time between donations
- **Average donation per year**



The average time between donations is 466 days

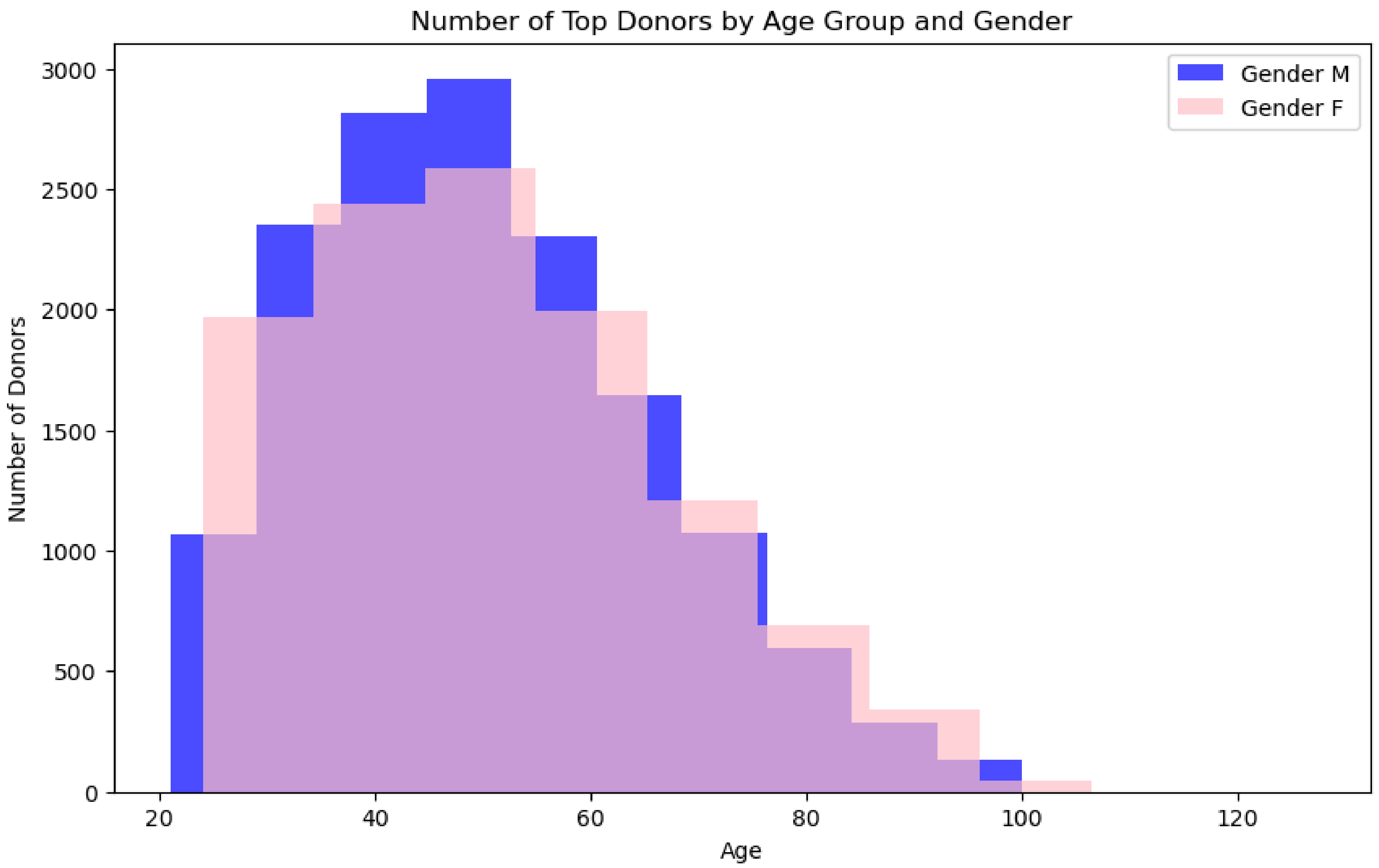
Historical Data

- Most **Frequent Month** of Donation
- **Dummy (Seasonality)** 1 if the campaign month matches the most frequent donation month for the donor; 0 otherwise.

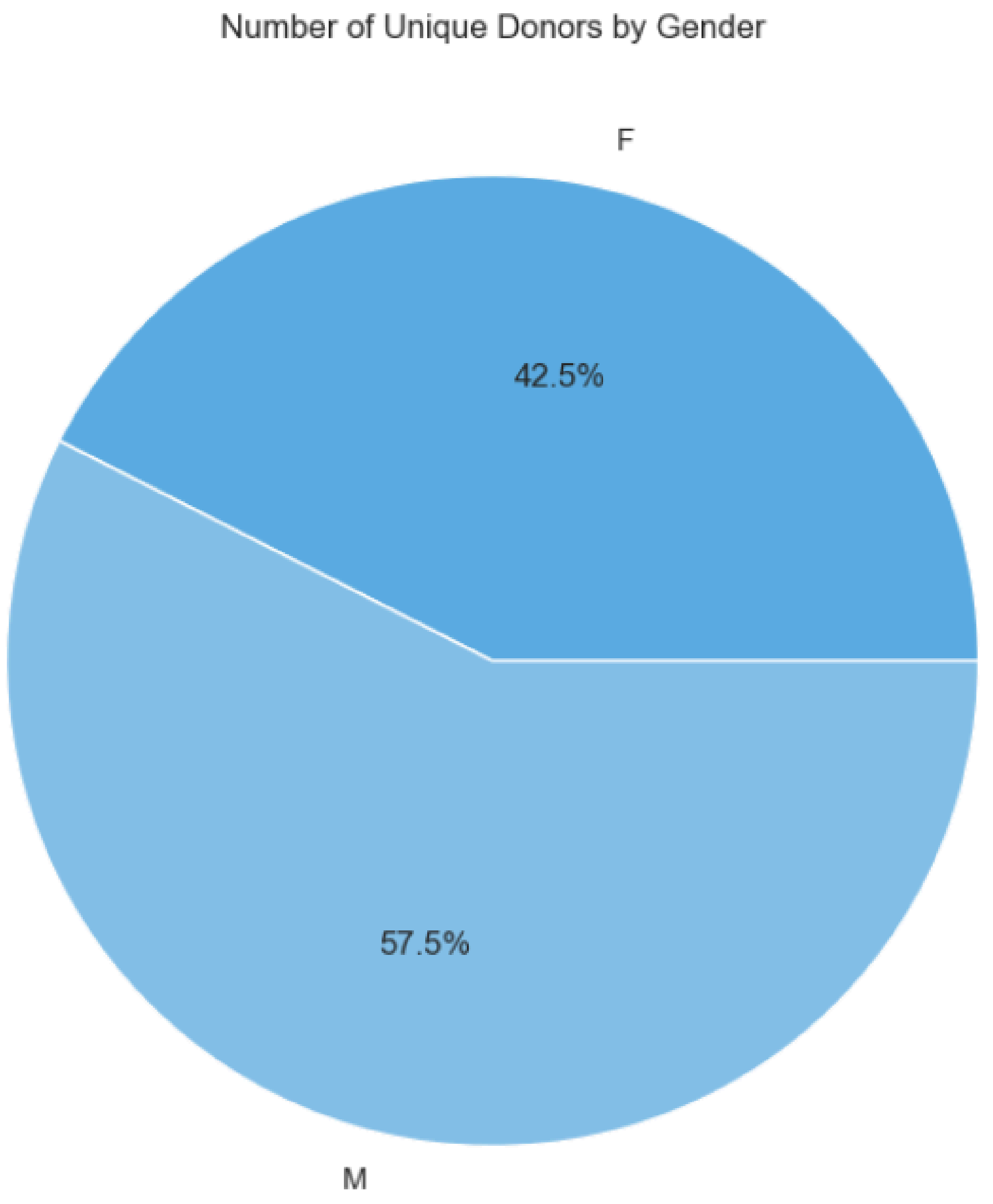


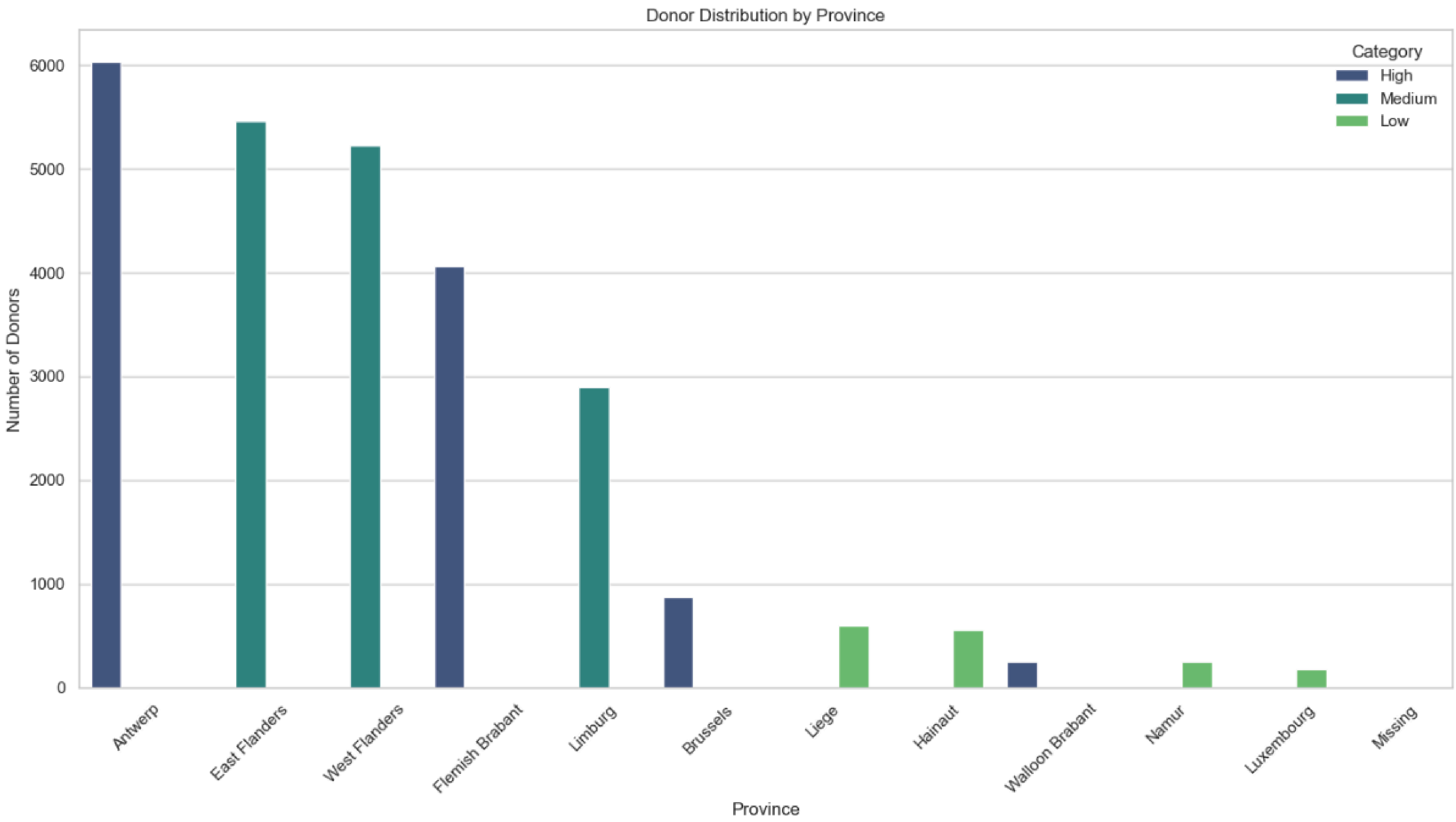
Demographic Data

Age
Gender
Region
Language
Level of Richness by Province (GDP Per capita)



More into Detail...





THE MODELS

Use of **hyperparameter** tuning to **improve performance**, and feature selection.

MODEL	Usage
Decision Trees	splits data into decision nodes
Logistic Regression	binary classification that predicts probabilities
Random Forest	ensemble model of decision trees to reducing overfitting
Gradient Boosting	corrects errors made in decision trees
K-Nearest Neighbors	predicts a label based on majority vote of nearest neighbors

HOW DO WE CHOOSE THE MOST EFFECTIVE ONE?

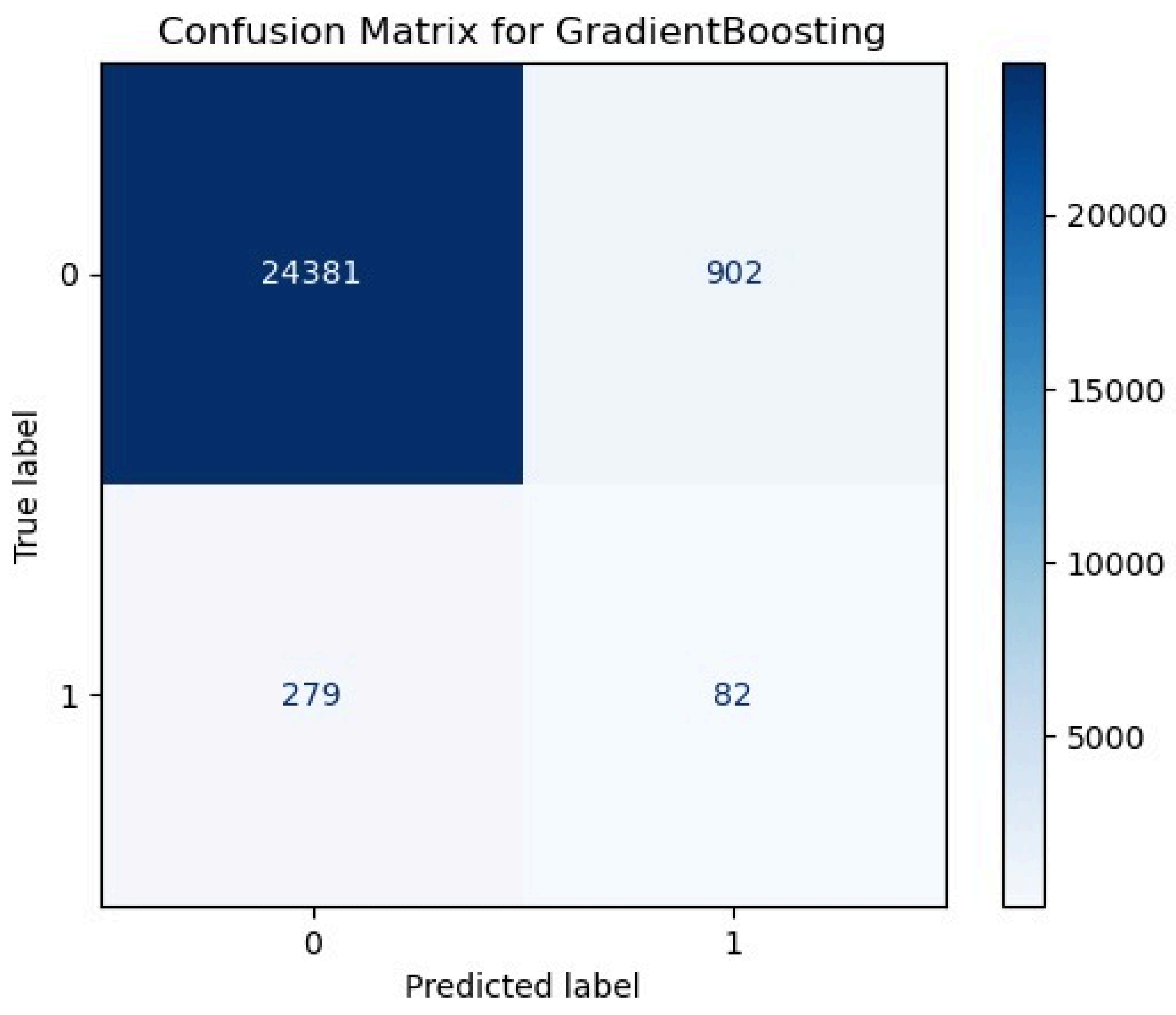
AUC

True positive rate

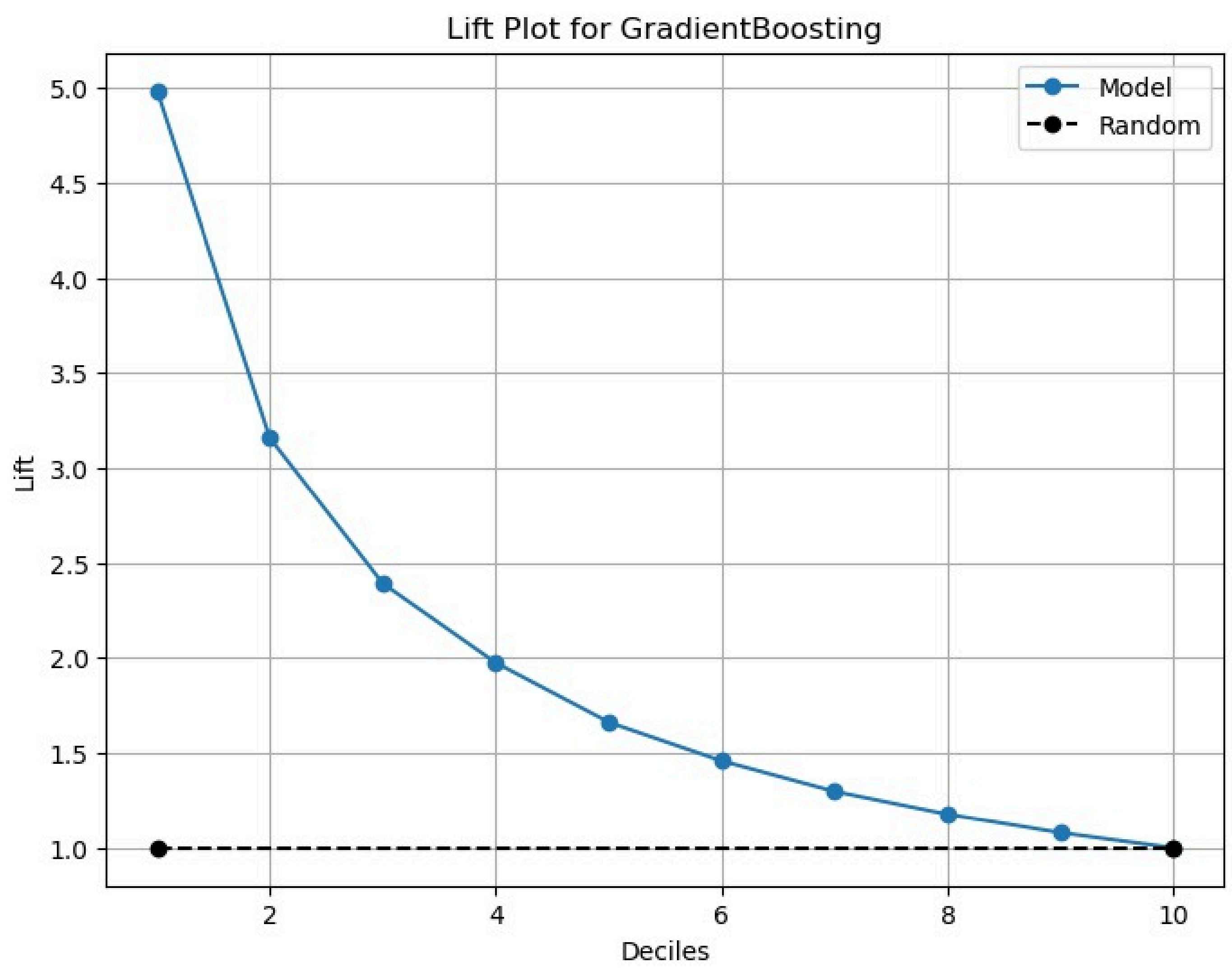
	Models				
	Random Forest	LG	DT	GB	KNN
AUC Test	73%	63%	73%	77%	59%
Acurancy Test	99%	99%	96%	95%	99%
TP	9	0	54	80	1
TN	25,268	25,282	24,638	24,386	25,279
FP	15	1	645	897	4
FN	352	361	307	281	360
%TP	2%	0%	15%	**22%**	0%
%FP	0%	0%	3%	**4%**	0%

OUR MODEL : GRADIENT BOOSTING

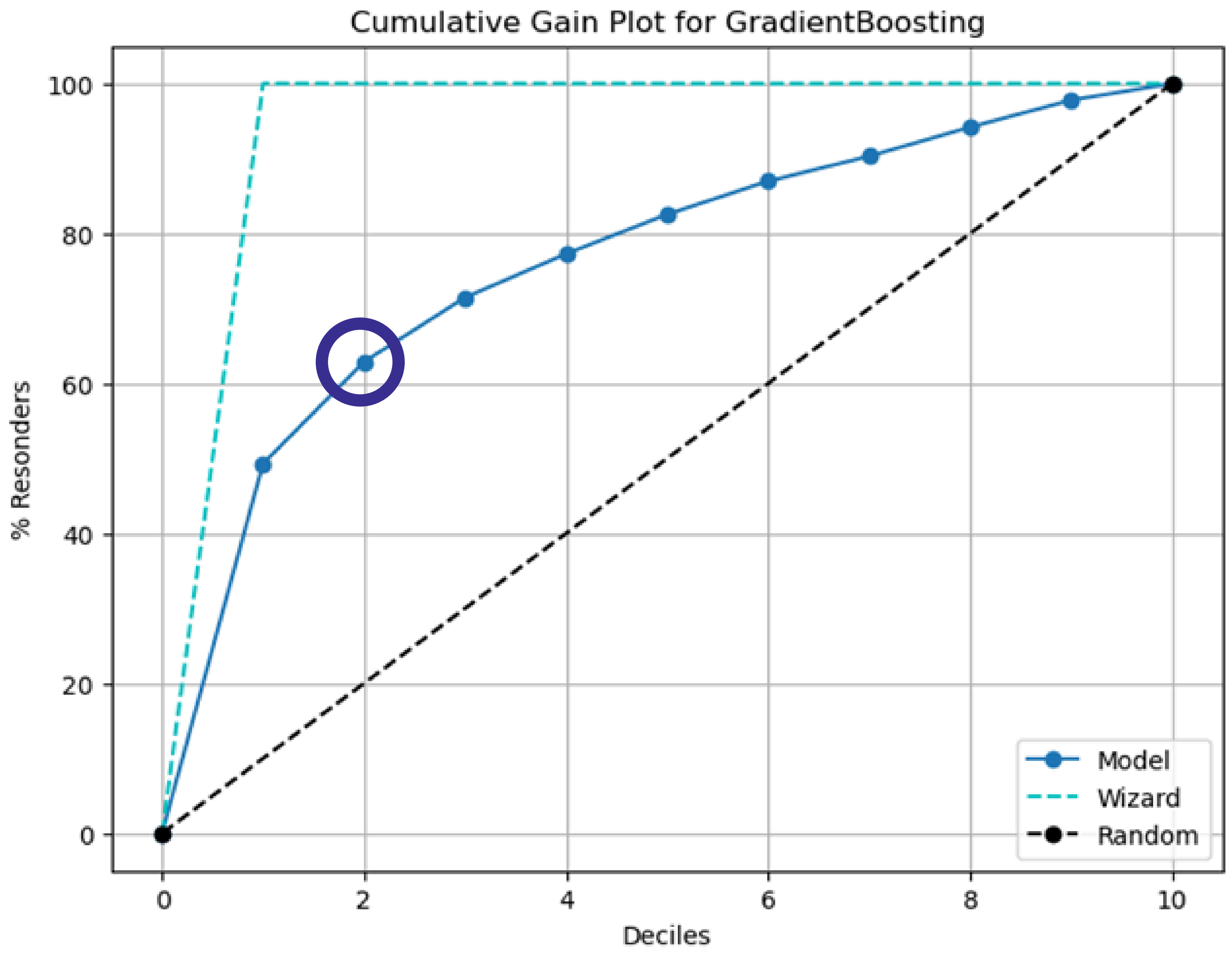
Confusion Matrix



Lift Chart



BUSINESS IMPLICATION: COST ANALYSIS



Cumulative Gains Chart

TOTAL number of **NEW DONORS** in new campaign = 26.522
Cost per letter 0.80 €

By **focusing on the top 20% of donors**, the company can significantly **lower costs** and **optimize donation outcomes**.

OUR MODEL VS RANDOM SELECTION

Metric	Random Selection	Gradient Boosting Model
# Possible Donors	26.522	26.522
Target Donors (20%)	5.304	5.304
Number of Donors Who Donated	95	**296 (62%)**
Total Donations Collected* (€)	€2.387	€7.400
Campaign Cost (€)	€4.244	€4.244
Profit projection (€)	**€-1.857**	**€3.156**

*Average Donation per donor €25.

RECOMMENDATIONS

Why our model should be used for future campaigns?

1.IMPROVED TARGETING:
The model focuses efforts on the 20% of most likely donors, leading to **higher response rates**.

2. COST SAVING:
The model reduces costs by **optimizing the number of contacts** sent compared to random selection.

3. HIGHER REVENUE:
Predictive optimization leads to better use of campaign funds, resulting in a **€3.156 profit projection**.

THANK YOU!