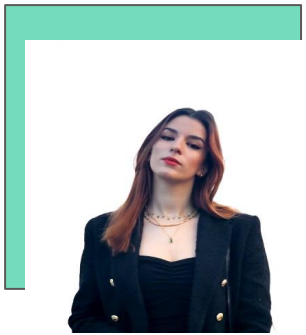
Biokind Analytics

ALS Texas

Hannah Choi, Sarah Lee, Frank Liu, Navya Bhagat, Christina Tzavara

Biokind Analytics

ALS Texas – Introduction



Christina Tzavara
Northwestern University
2025

Frank Liu
New York University
2025

Hannah Choi
Duke University
2026

Sarah Lee
University of Chicago
2025

Navya Bhagat
University of Rochester
2025

The ALS Association of Texas

Mission: “to discover treatments and a cure for ALS, and to serve, advocate for, and empower people affected by ALS to live their lives to the fullest”

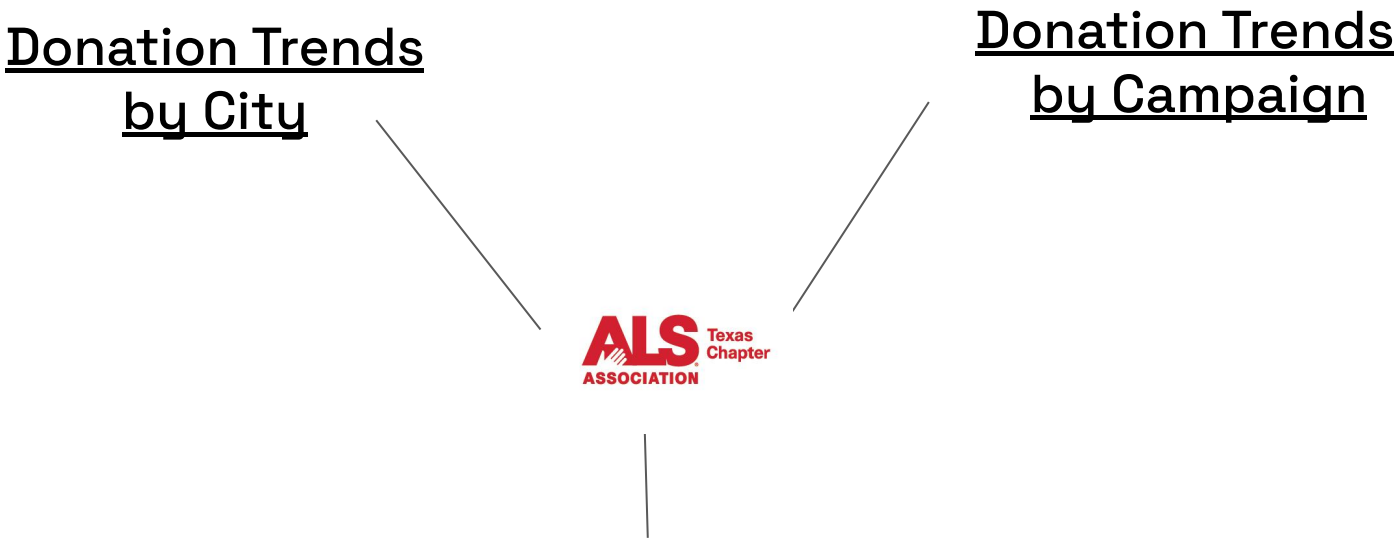
Our partner: Kathryn Berkson-Sheppard
Managing Director of Development

ALS Texas Chapter
ALS ASSOCIATION

https://alstexas.org/about-us/

Biokind Analytics

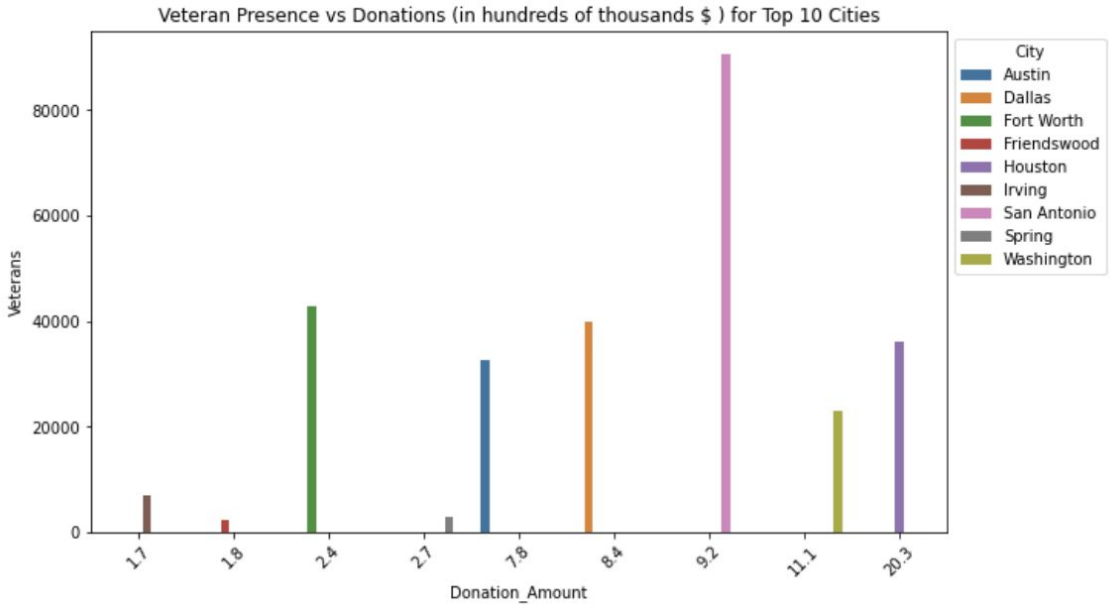
ALS Texas - Areas of Focus



The cities donating the most…

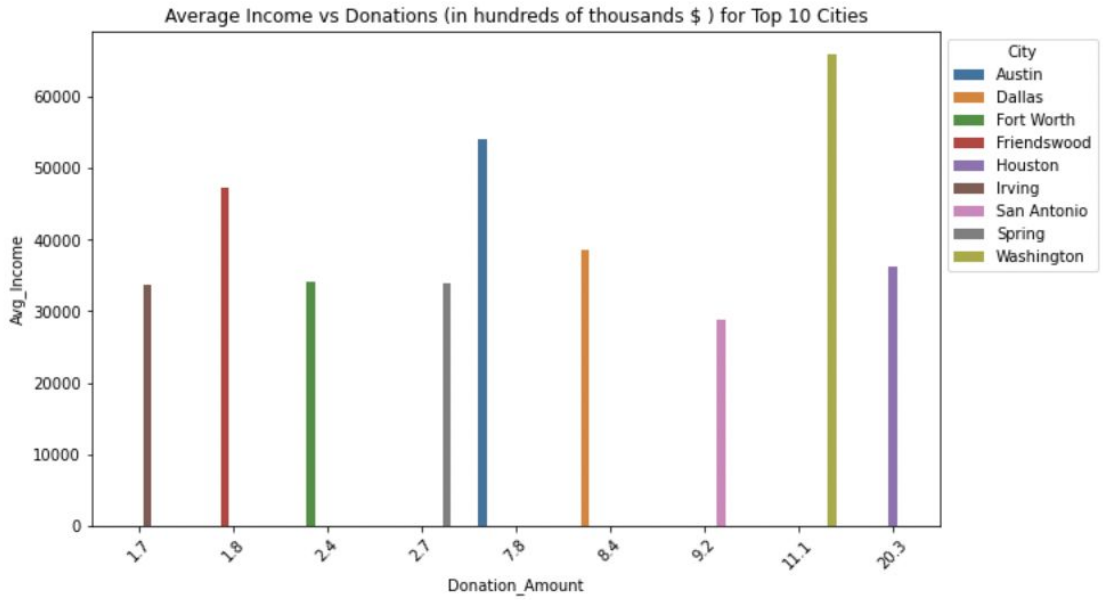
City	Donation Counts	Total Donations	Average Donation
Houston	4095	2032153.26	496.252322
Washington	63	1113701.11	17677.795397
San Antonio	3755	920452.24	245.127095
Dallas	2540	844373.08	332.430346
Austin	3421	776514.91	226.984773
Grapevine	80	368257.67	4603.220875
Spring	876	272511.37	311.086039
Fort Worth	928	240118.19	258.748050
Friendswood	155	178407.85	1151.018387
Irving	190	168359.13	886.100684

On veteran presence in the top 10 cities…



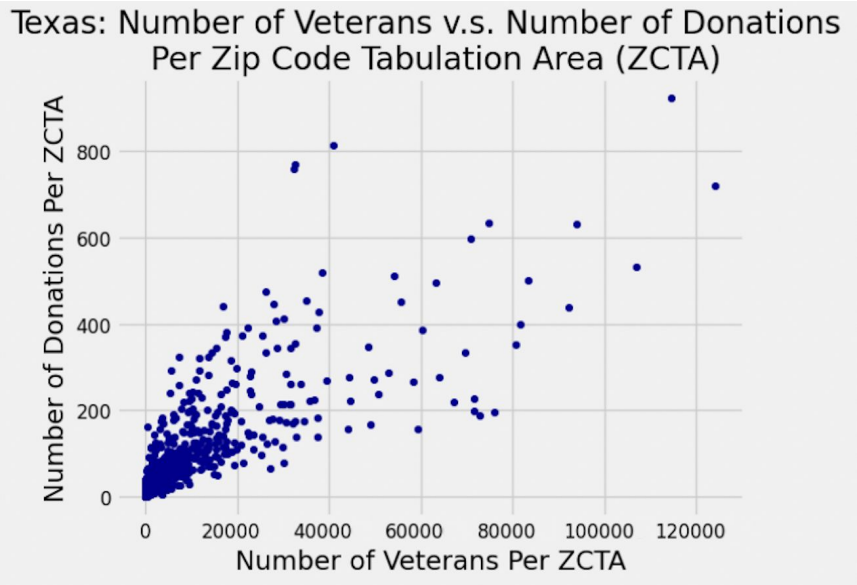
Mean Veteran Population: 30.795

On the average income per capita in the top 10 cities…



Mean Average Income per Capita: $41.340

Relationship between In-State Veteran Presence & Donations

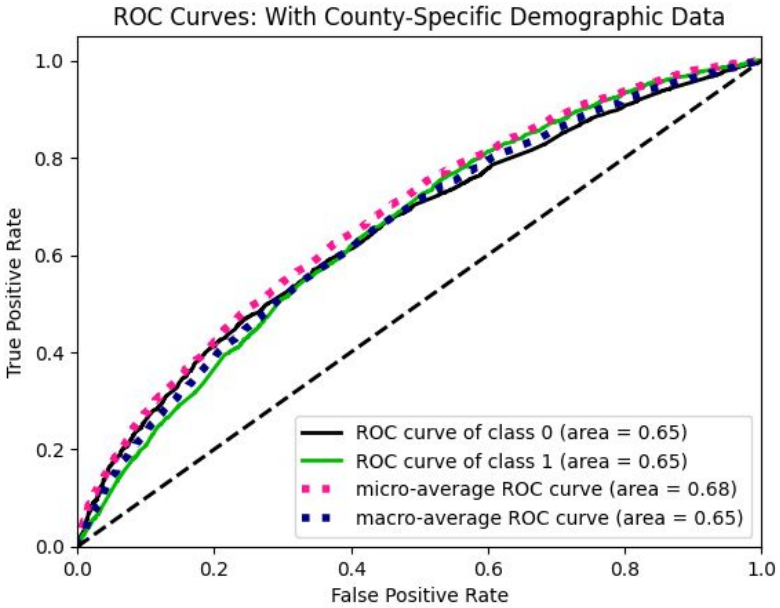


A positive relationship exists between the number of veterans and the number of donations per county

<u>Top 3 Counties Associated with ZCTAs with Largest Veteran Presence:</u>

Harris County, Houston
Bexar County, San Antonio
Tarrant County, Fort Worth

Logistic Regression Model Using County-Specific Data



Incorporating **county-specific demographic data** improves the logistic regression model's ability to predict **donor retention status** accurately **from ~56% to 65%**

Strongest county-specific predictors for donor retention include:

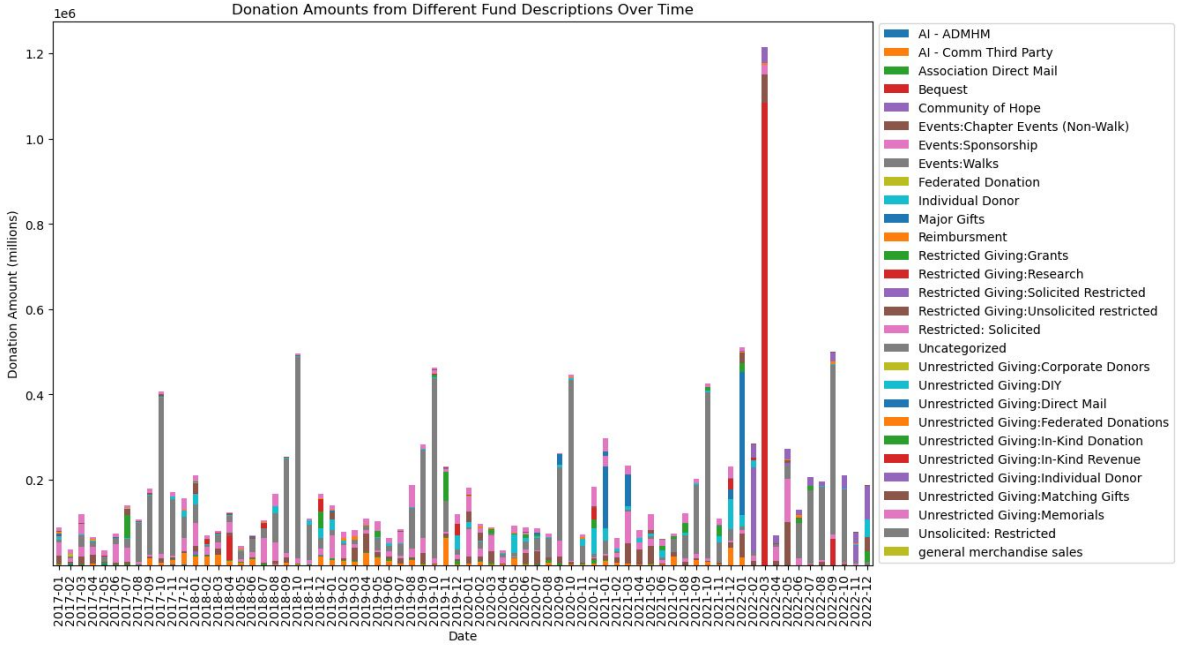
- Number of married individuals & families
- Number of veterans
- Number of nonprofits
- Population

Donation Trends by Campaign

Statistics of the top 10 fund descriptions with the highest total donations

Fund Description	Minimum	Median	Maximum	Count	Sum	Mean
Events:Walks	1.0	50.0	50000.0	43400	5410820.0	124.67
Events:Sponsorship	25.0	2500.0	50000.0	384	1256504.0	3272.14
Bequest	57000.0	60674.0	1083782.0	3	1201456.0	400485.33
Events:Chapter Events (Non-Walk)	0.0	100.0	25000.0	2145	901038.0	420.06
Unrestricted Giving:Memorials	4.0	100.0	12000.0	4535	672182.0	148.22
Major Gifts	5000.0	23750.0	250000.0	12	603567.0	50297.25
Individual Donor	0.0	100.0	15000.0	2144	584293.0	272.52
AI - Comm Third Party	0.0	87.0	30042.0	1100	515136.0	468.30
Restricted Giving:Grants	300.0	10000.0	55000.0	27	370866.0	13735.77
Unrestricted Giving:Individual Donor	0.0	100.0	10000.0	1332	312690.0	234.75

Donation Trends by Campaign



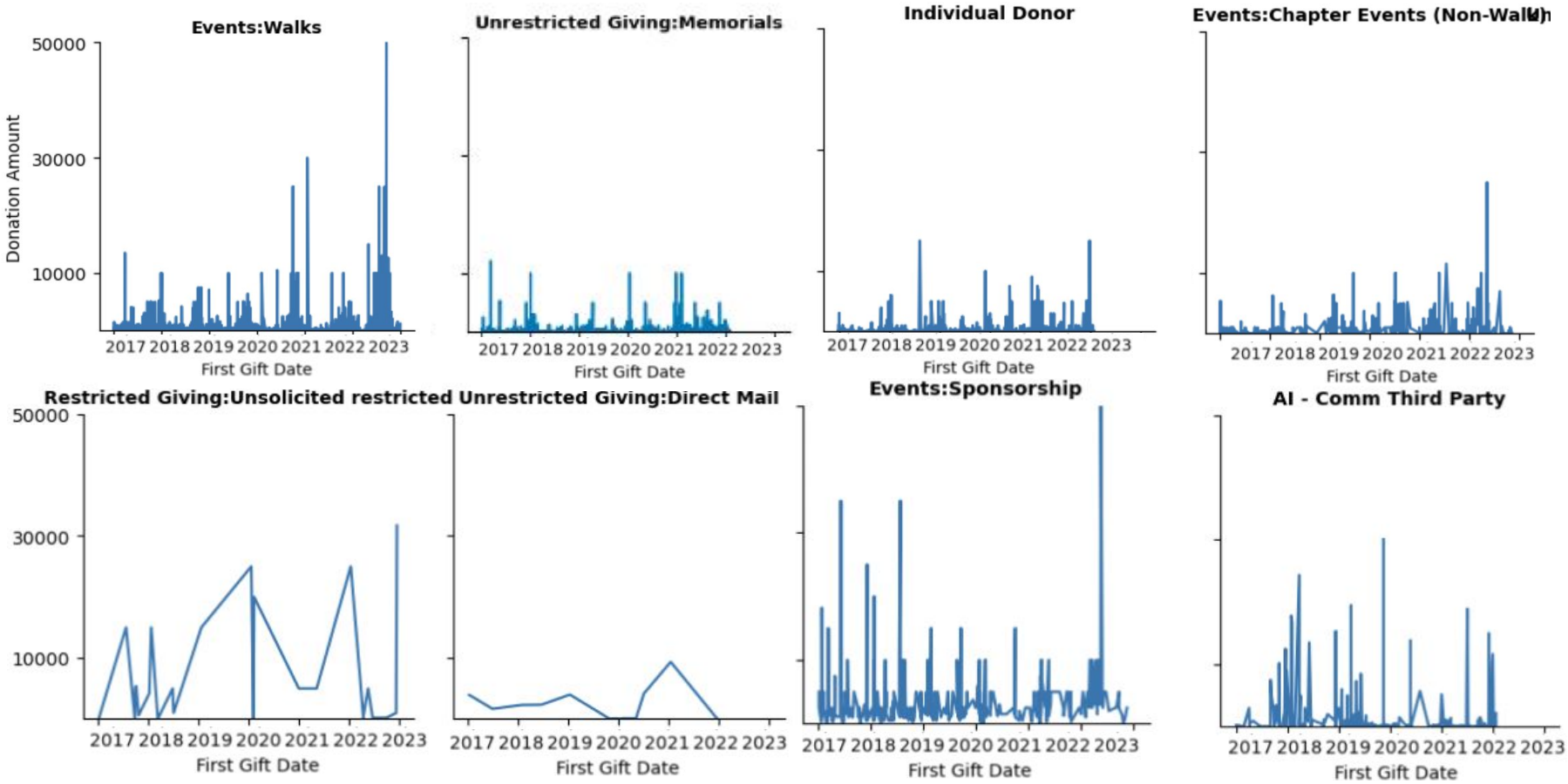
Donations peak every year in October (from Walks, Restricted & Uncategorized events).

Exception:

- Jan 2022
  - Spike in AI-ADMHM, **Major gifts**, Direct Mail
- March 2022
  - Spike in Bequest, **Research, In-Kind Revenue**

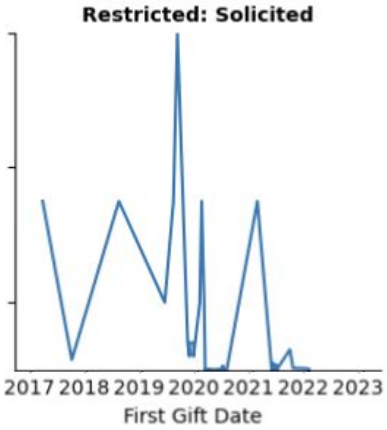
Biokind Analytics

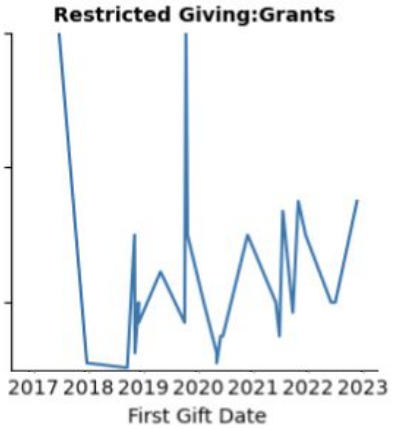
ALS Texas – Analysis

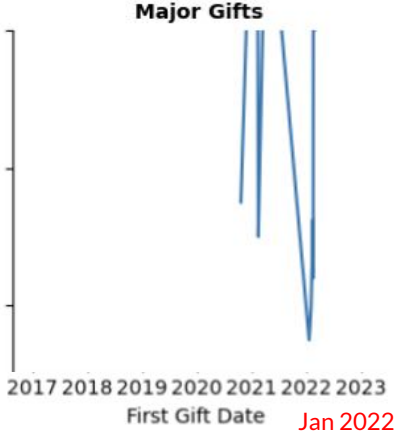


Biokind Analytics

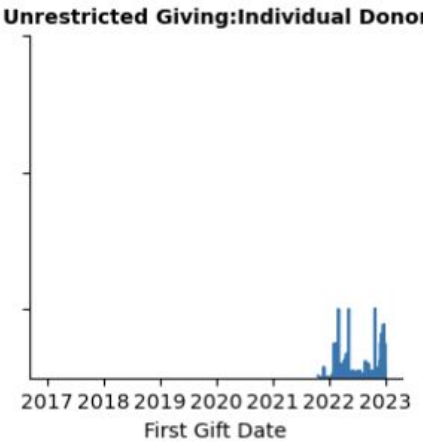
ALS Texas – Analysis

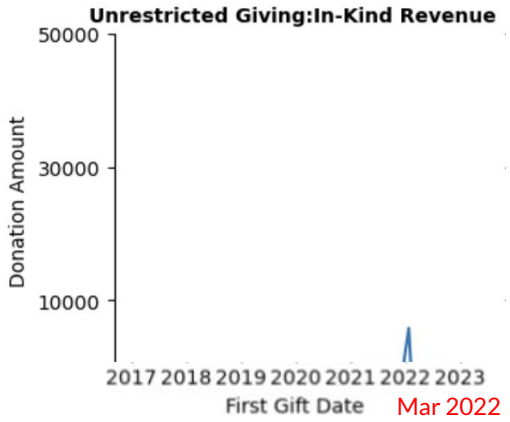




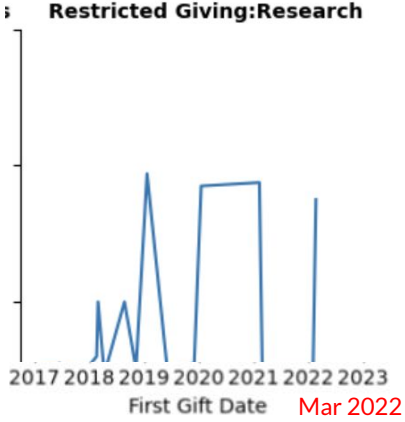


Jan 2022



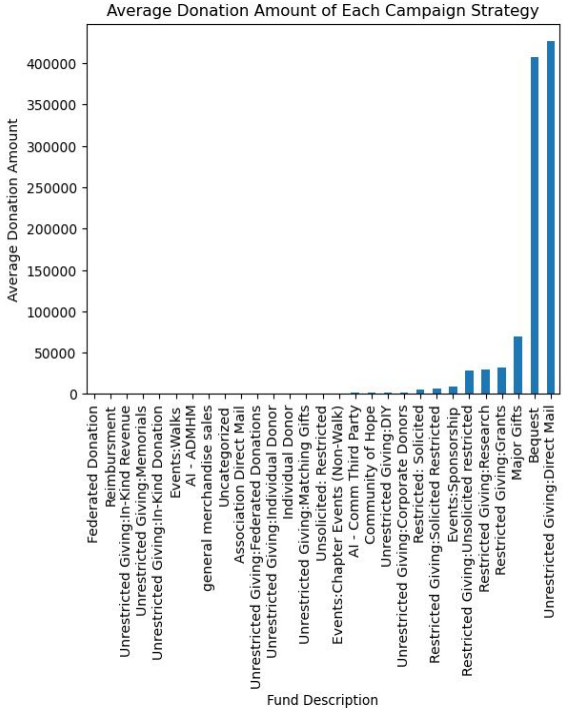


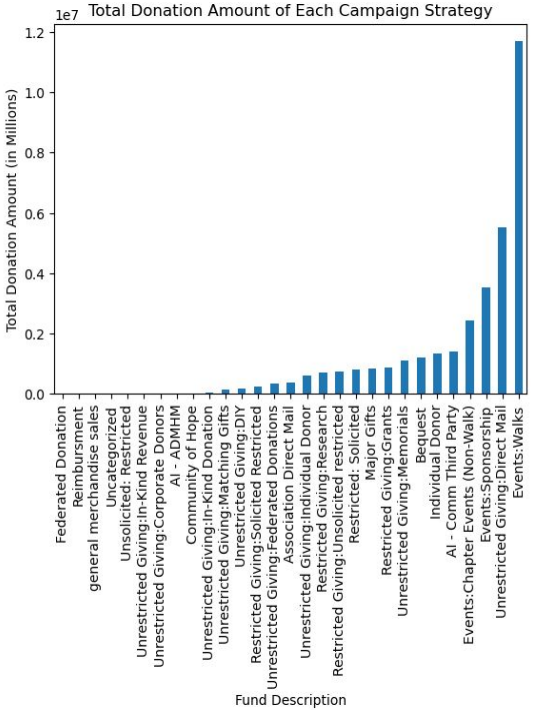
Mar 2022

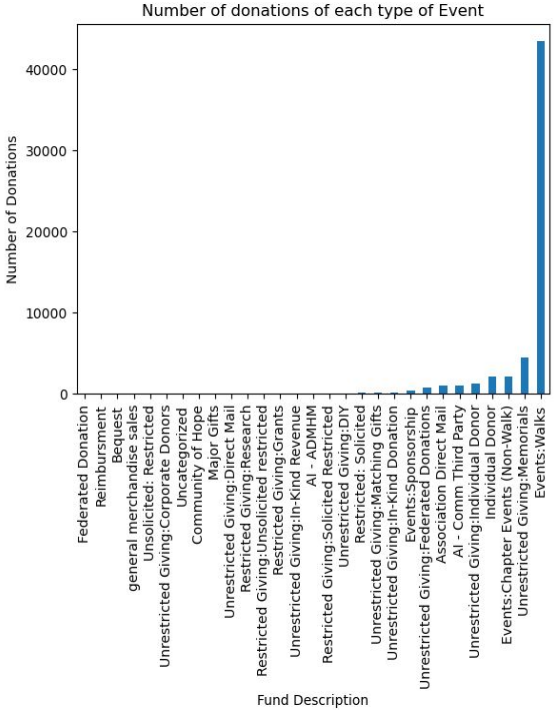


Mar 2022

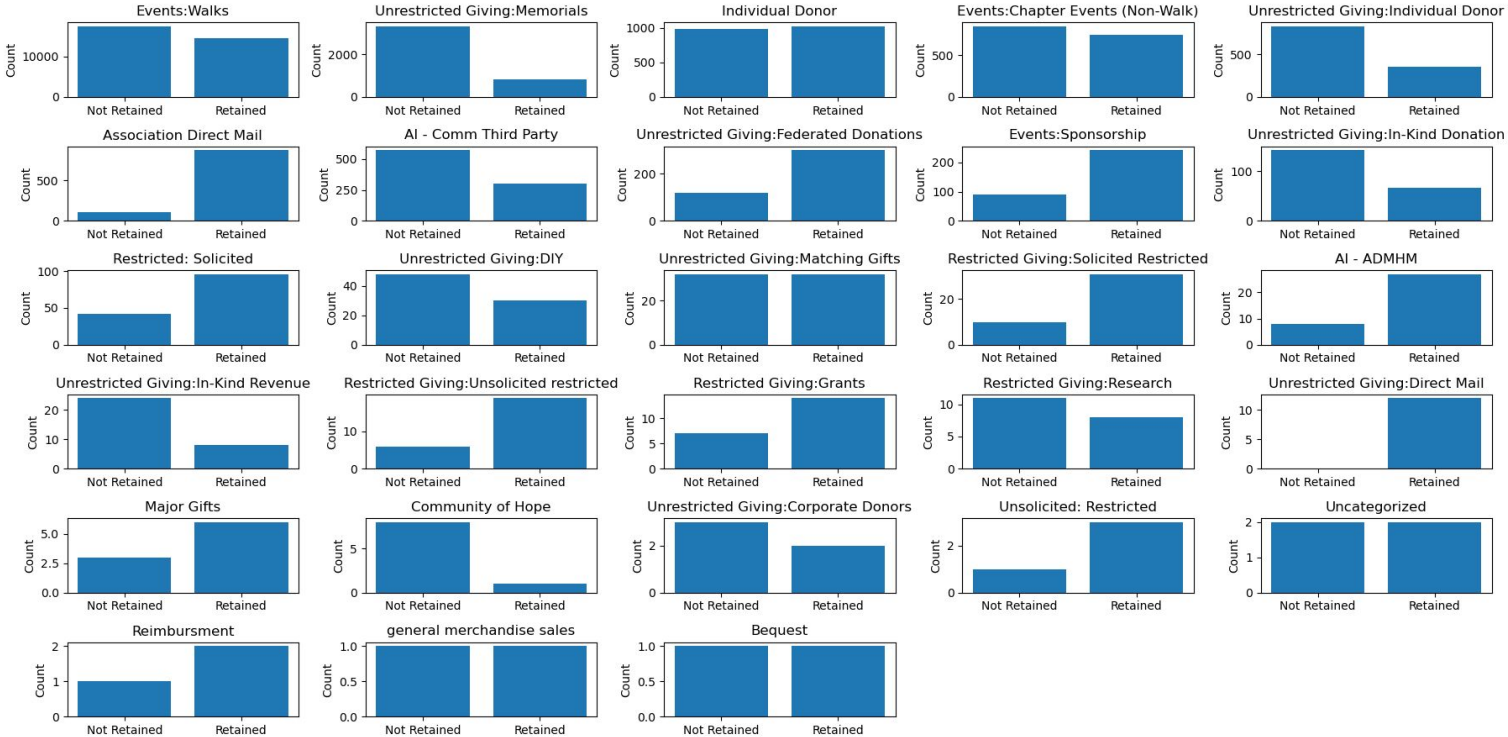
Visualizations of different Campaign Strategies







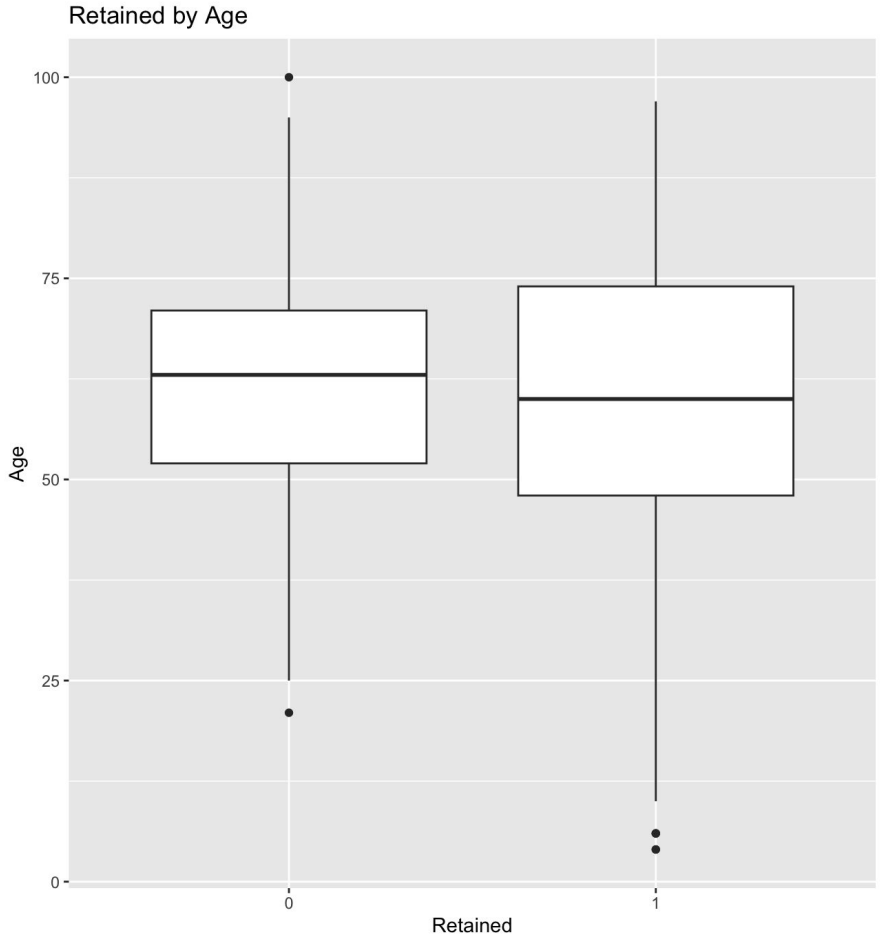
Donor Retention by Event Type



Retained Donor by Age

Retained median age: **60**

Not retained median age: **63**

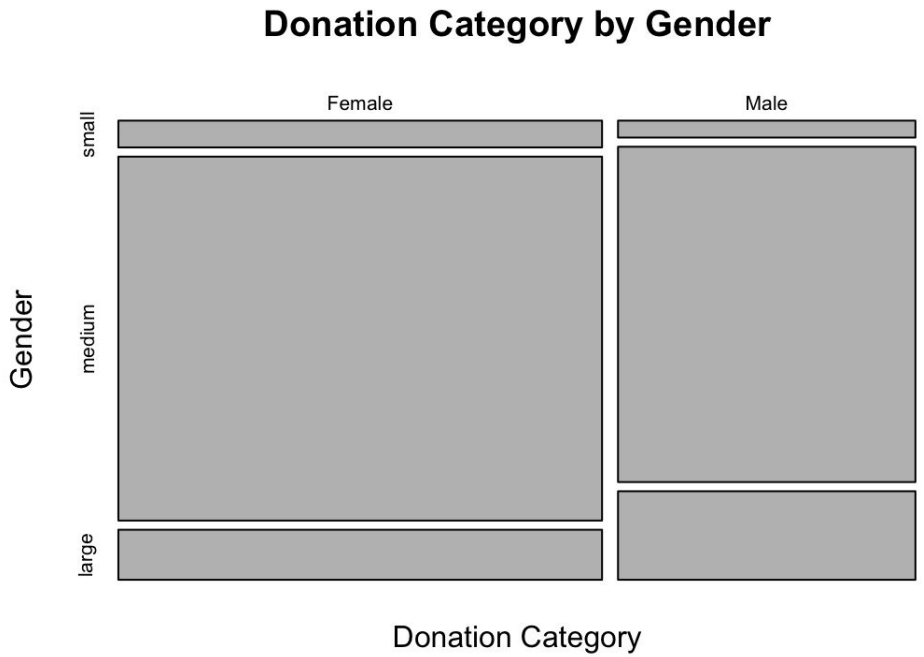


Donation Amount by Gender

Small: Bottom 10% of donors
$0-$20

Medium: Middle 80% of donors
$20-$250

Large: Top 10% of donors
$250-$1083782

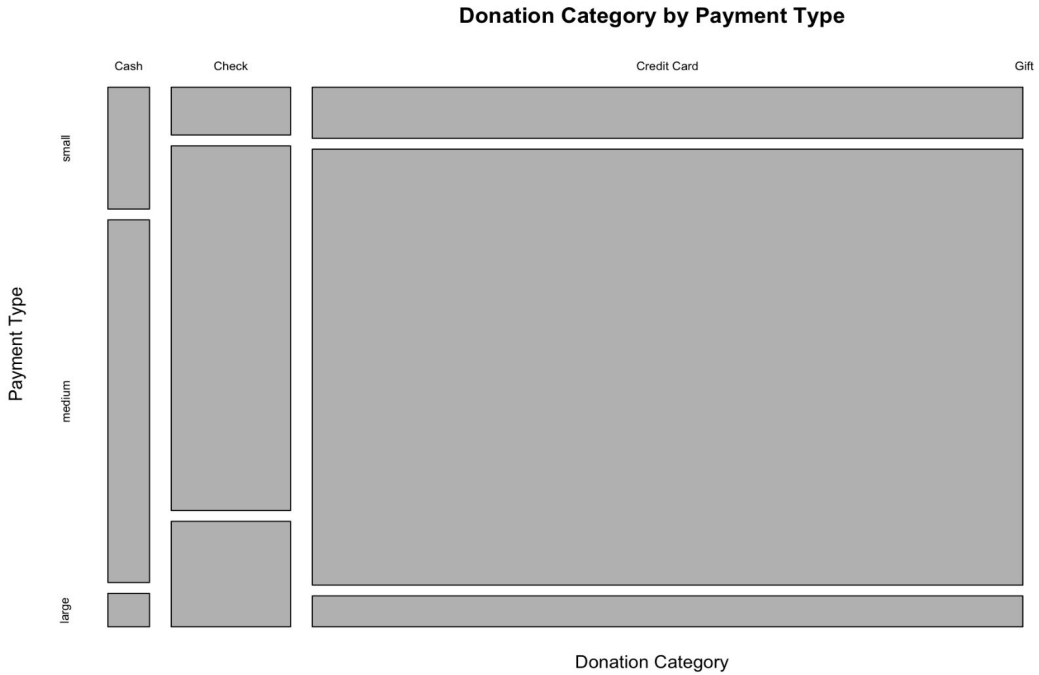


Donation Amount by Payment Type

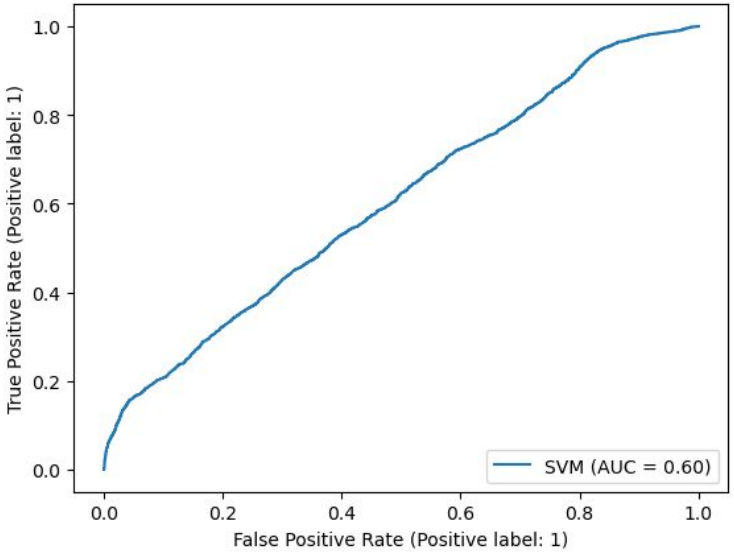
Small: Bottom 10% of donors
$0-$20

Medium: Middle 80% of donors
$20-$250

Large: Top 10% of donors
$250-$1083782



Prediction Models



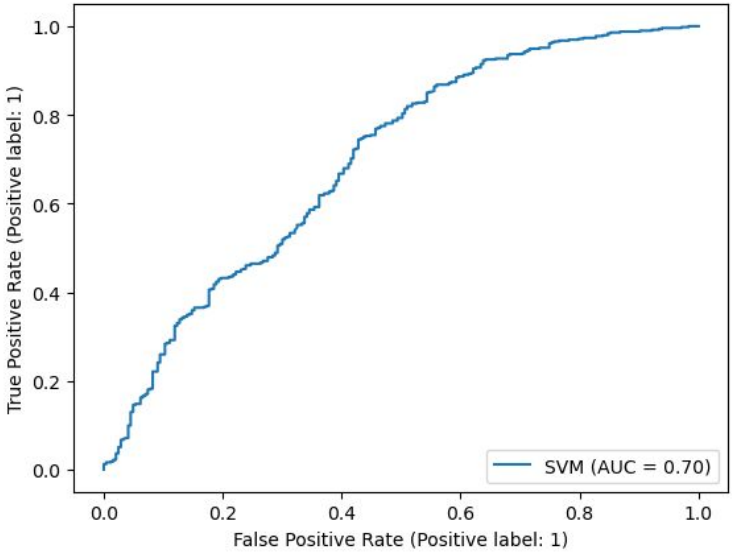
Performance Metrics

```
SVM model performance
Confusion Matrix:
 [[4721  251]
 [3268  636]]
Accuracy: 0.6035376295628662
Precision: 0.6539806072131593
Recall: 0.5562135664640017
F1 Score: 0.4969954901069472
ROC-AUC: 0.6047448446958708
```

Best Predictors

```
===Lowest Performance Scores of each Metric of SVM model===
Lowest AUC: Direct Mail
Lowest Precision: Direct Mail
Lowest Accuracy: Gender
Lowest Recall: Gender
Lowest F1: Gender
```

Prediction Models (Without Missing Gender)



Performance Metrics

```
SVM model performance
Confusion Matrix:
 [[ 69 174]
 [ 26 467]]
Accuracy: 0.7282608695652174
Precision: 0.7274324657196815
Recall: 0.6156061402849773
F1 Score: 0.6159585903175646
ROC-AUC: 0.6969590731141329
```

Best Predictors

```
===Lowest Performance Scores of each Metric of SVM model===
Lowest AUC: Direct Mail
Lowest Precision: Memorials
Lowest Accuracy: Walk
Lowest Recall: Walk
Lowest F1: Walk
```

Biokind Analytics

ALS Texas – Analysis

Prediction Models Usage:

Model: SVM (Trained with no Missing Gender)

Input:

	Donation Amount	Gender	lat	lng	population	density	Walk	Memorials	Chapter	Individual Donor	Direct Mail	Comm Third Party	Federated Donations	Sponsorship	Quarter
36202	52.5	1	32.88227	-97.34895	47569.0	1182.0	0	1	0	0	0	0	0	0	3
19057	10.0	1	30.55851	-97.91473	81311.0	252.4	0	0	0	0	1	0	0	0	3
30761	10.0	-1	32.77943	-97.51811	44125.0	395.3	1	0	0	0	0	0	0	0	4
1695	20.0	-1	29.72477	-95.36498	34197.0	2519.3	1	0	0	0	0	0	0	0	3
43004	30.0	1	32.26006	-99.52872	7673.0	11.6	0	1	0	0	0	0	0	0	3
958	100.0	1	29.57584	-95.13406	24761.0	1644.6	1	0	0	0	0	0	0	0	4
36557	100.0	1	33.57235	-96.17730	14726.0	40.4	0	0	0	0	0	1	0	0	4
22842	100.0	-1	30.05392	-95.15884	27274.0	1256.1	1	0	0	0	0	0	0	0	3
33060	50.0	1	31.85005	-106.45417	32148.0	1038.6	1	0	0	0	0	0	0	0	4
43449	500.0	-1	32.32614	-100.92807	7758.0	5.0	0	1	0	0	0	0	0	0	4

Output:

Actual:

	Predicted Not Retained Probability	Predicted Retained Probability	Retained
36202	0.676308	0.323692	0
19057	0.161712	0.838288	1
30761	0.243525	0.756475	0
1695	0.233041	0.766959	0
43004	0.721923	0.278077	0
958	0.245532	0.754468	1
36557	0.560112	0.439888	0
22842	0.237782	0.762218	0
33060	0.244651	0.755349	1
43449	0.727194	0.272806	0

Data-driven recommendations

- When it comes to increased veteran presence, ALS should focus on Harris County, Bexar County, and Tarrant County in Texas
- ALS Texas should target the $35-65k income range since it prevails in the cities donating the most
- Direct mail is one of the best methods of retaining donors
- Collecting more quantitative data (e.g. Age/Gender) can improve donor retention prediction models
- Employ county-specific demographic datasets to enhance donor retention prediction (e.g. number of married individuals, number of veterans)