CIS/SCM 593 Project Report

Team 010	Darshan Kamalakar, Austin Johnson, Julis Huang, Zih Han Shen
Topic	Current and Past Donor Data Analysis
Client	Grand Canyon Council, Scouting America

Under the guidance of
Joseph Cazier, Ph.D and CAP (Certified Analytics Professional)
Clinical Professor,
Department of Information Systems

Contents

1. The Problem

The Boy Scouts of America (BSA) was founded in 1910 with the mission statement of "To teach [boys] patriotism, courage, self-reliance, and kindred values." They successfully grew over the years while maintaining similar values throughout and celebrated a golden age starting around the 1960s to 1972 when it reached its peak membership of over 6.5 million nationwide and has been slowly declining since. Having operated for over 100 years, there have been countless hurdles to overcome as well as hardships. In more recent times with sexual assault allegations, (leading to filling for bankruptcy), the Covid-19 pandemic, rebranding and policy changes within the scouts and lastly, a new technological era.

We believe that given the scouts values and changing times, that rebuilding and expanding the scouts starts at the foundation and state level. In order to most effectively do this, they need the support from the community financially as well as with volunteers. Our goal is to identify and breakdown segments within the community to better target, engage and interact with them to rebuild and reach these goals.

Scouting Arizona, Grand Canyon Council in recent years has been facing a stagnation in donations and fundraising, with a need to reverse these trends to rebuild and strengthen the organization. The current growth does not align with the strategic goal of a 10% annual growth of donors and overall fundraising. To drive sustainable growth, Scouting Arizona, Grand Canyon Council needs to focus on more effectively and efficiently targeting its donors and donations. By analyzing data, different market segmentation can be formed on the basis of geographic location, age group(s) and donor consistency.

Based on the market segmentation, tailored engagement strategies will be developed to ensure effective communication and maximize donation potential. The goal is to raise growth of donors and overall fundraising by the range of 3% to 5% annually, in order to support the scouts and all the projects they do within the Arizona community. These efforts will help strengthen the core values of Scouting Arizona Grand Canyon Council of, communication and leadership, cultivating social capital and achieving goals while being accountable.

The action plan will be crafted based on this focus, leveraging data and analysis to refine strategies and address the specific needs of key demographics.

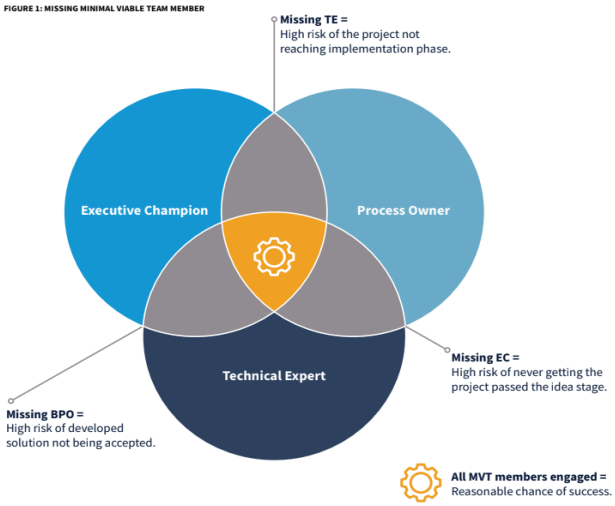
- The problem to be solved is reversing the stagnation of growth within the Scouting Arizona Grand Canyon Council by increasing the overall number of donors, donations and fundraising. By analyzing data and creating market segmentations, Scouting Arizona can then more efficiently and effectively tailor engagement strategies to each segment and maximize donation potential. By doing so, Scouting Arizona will be able to meet this 2.5% annual growth for donors, donations and fundraising.
- We have pulled from the 2020-2024 Impact reports that membership and fundraising/donations are stagnant. We also saw within their strategic planning, a desire to "Grow the number of donors and overall fundraising by 10%, annually." Donations nor membership do not match with this strategic planning currently.
- We have maintained the overall goal for growth but have set a 2.5% annual growth change. We believe the 2.5% annual growth change more aligns with what can reasonably be expected given the data at hand. As further analysis is conducted into each market segment and how effective each may be, this could change.

- Our problem statement and our goals are clearly time bound on a year-to-year basis to see the desired growth and impact. This yearly analysis aligns with our clients strategic plans and also allows for not too much or too little time to evaluate the effectiveness and efficiency of our actions.
- This problem statement was designed to specifically align with the clients strategic planning while also maintaining the values, boundaries and goals of our client.
- In order to fulfill the overall goal of increasing donations and fundraising, analysis was conducted and an initial briefing was held. From the initial briefing, the scale of aspirations as well as the direction for the project were modified to better fit and reflect our clients goals.
- Our team understands the importance of being flexible and understands goals and that situations can change. Additionally, the team is aware that just because you “come up with a solution” this solution may be wrong or not work.
- Because of this, the team has approached and structured the problem in a way that is critically not too narrow or too broad with its scope. By doing so, the team has allowed itself to pull and conduct analysis on our data to fulfill the problem statement while also being able use this information as paths to other solutions, should the main one be insufficient.
- Our problem statement and plan does allow for the problem to be solved and applied at the highest level possible. By designing the problem statement and solution in terms of market segmentation, not only does this allow for this lower level solution to be solved within Arizona, but it specifically tailors solving this similar problem across each and every state within the overarching organization of Scouting America.

The solution using our problem statement very specifically adds value by breaking down the market into segmentations based on previous donation records. In the long run this will firstly, save money going out of the organization by not blanket advertising to all groups and everyone possible. Secondly, this approach will allow for more specific and targeted advertising which in turn will result in higher engagement and maximize donation potential. In order for this solution to work most effectively, this will require action from our client to coordinate with a donor relations and/or marketing team to advertise and engage with each market segment, thus resulting in a more personalized outreach. By taking these actions, we will then be able to compare past and present KPIs of “number of contributions”, "amount of contributions” and “average contribution” for each market segment. This will give us even more insight and feedback on how well our target advertising has been to each market segment and will even give further feedback and insight by having control groups based on segments that were not targeted and advertised to.

2. Team

This revision aligns with the MVT framework, ensuring that you have an **Executive Champion**, **Business Process Owner**, and **Technical Experts**, while also recognizing additional stakeholders needed for project success



Resource:
Beyond AI: The Minimal Viable Team Needed for Successful Business Transformation with AI and Analytics
(Dr. Joseph Cazier, Dr. Rudi Pleines)

Stakeholder	**Role**	**Expectations from the team**	**What the team needs**
Andy Price (Grand Canyon Council)	Executive Champion (Client PoC)	A solution that increases donor engagement and fundraising impact. Regular updates & key insights for decision-making.	Clear feedback on fundraising priorities & donor segmentation. Access to historical donor data and insights.
Marketing Team (Grand Canyon Council)	Process Owner	Actionable insights on donor engagement to enhance campaign effectiveness.	Collaboration on donor segmentation strategies. Alignment with ongoing marketing initiatives.
Prof. Joseph Cazier (ASU)	Faculty Advisor	A structured approach with clear milestones. Demonstration of analytics best practices.	Guidance on methodologies, project structuring, and feedback on implementation.
MSBA Team 010 (ASU)	Technical Experts (Data Analysts & Project Executors)	Hands-on experience applying analytics to a nonprofit use case.	Clear project scope, data access, defined success metrics, and collaboration with process owners and executive champions .

GCC Leadership (Grand Canyon Council)	Additional Executive Champions	Demonstrated ROI on analytics-driven donor strategies. Increased donor retention & engagement.	Buy-in for implementation, access to donor relationship history, and alignment with organizational goals.
IT/Data Team (Grand Canyon Council)	Technical Support	Secure and efficient integration of analytics tools.	Access to relevant donor data, compliance with data privacy policies, and technical infrastructure support.

3. The Data

Data Overview and Relevance to the Problem

Data is at the core of this project and is essential to its success. However, data is only valuable if it effectively addresses the problem at hand. This section outlines the datasets available, verifies their relevance and quality, and explores additional data sources that may enhance our analysis.

The problem we aim to solve involves understanding donor behaviors, optimizing fundraising strategies, and improving donor retention. To achieve this, we will analyze **Gift Records** to identify patterns in giving, donation methods, and fund allocations. Additionally, we will leverage **Constituent Records** to assess donor demographics, engagement levels, and historical giving trends.

While we are in the early stages of data analysis, we have conducted an initial review of the available datasets to ensure they contain meaningful and actionable insights. This review has confirmed that the datasets include a substantial volume of records, covering various aspects of donor behavior. However, further data cleaning will be necessary to address potential issues such as missing values, inconsistencies, and data completeness.

Resource Description	Location	Access	Need	Constraints
Gift Record: Includes donation amount, type, fund, and payment method. (66,019 records, 14 variables).	Gift Record.xlsx	Access is available to authorized project team members	● Essential for analyzing donation trends, payment preferences, and fund allocations. ● Helps identify key donor behaviors and optimize fundraising strategies.	● The data is confidential and must comply with privacy and data protection regulations. ● Access should be restricted to authorized users only, and no external sharing of the data is allowed.

Resource Description	Location	Access	Need	Constraints
Constituent Records: Contains donor personal details, giving history, and engagement with the organization. (28,348 records, 32 variables).	Constituent Records.xlsx	Access is available to authorized project team members	● Critical for understanding donor demographics, engagement levels, and retention patterns. ● Helps in donor segmentation and targeted outreach.	● The data is confidential and must comply with privacy and data protection regulations. ● Access should be restricted to authorized users only, and no external sharing of the data is allowed.
Synthetic Data: Artificially generated data derived from donor records and behavior prediction models. Includes features such as donor propensity scores, predicted engagement, and normalized behavior segments**.**	DonorProfile.csv, ML-Powered_Propensity_to_Donate_Prediction, normalized_dataframe_segments.csv, Donation - AI Use Cases for Boy Scouts of America_Feb 2025	Access is available to authorized project team members	● Enables testing and development of ML models without exposing real donor data. ● Supports simulation of donation behavior and segmentation analysis.	● Though synthetic, data should still be handled responsibly to prevent misinterpretation. ● Access should be limited to the project team for intended analytical or modeling purposes.

Data Quality and Availability

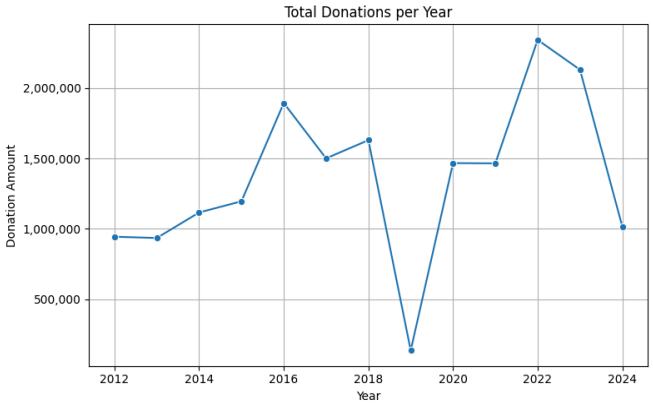
A preliminary review of the datasets indicates that they provide robust coverage of donor transactions and engagement. The **Gift Records dataset** spans multiple years and includes detailed financial transactions, which will be instrumental in identifying key giving patterns. The **Constituent Records dataset** provides donor demographic information and engagement history, which are essential for segmentation and predictive modeling.

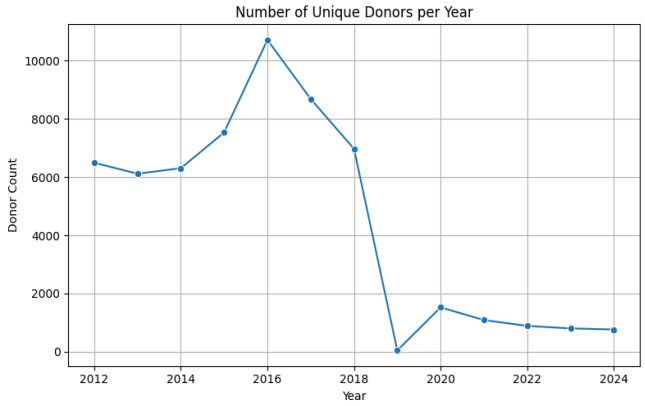
While the datasets appear comprehensive, we anticipate challenges related to **data completeness, accuracy, and standardization**. For instance, there may be missing or inconsistent records in donor engagement fields. As a result, an initial data cleaning and validation phase will be necessary to ensure the reliability of our findings.

Data Trends and Key Observations

Building on our initial data assessment, we have conducted a preliminary analysis of donation trends, membership fluctuations, and external influences affecting fundraising performance. The key observations are:

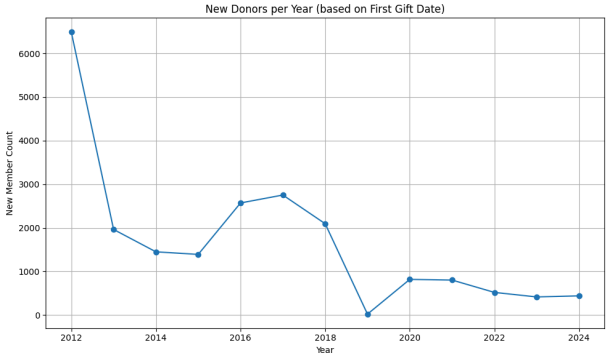
1. Donation Trends (2012–2024)





- Historically, 70-80% of donations came from individuals until 2019, after which the proportion fluctuated between 35-60%, reaching a record low of 10% in 2024.
- One-time donations dominated until 2018, with "other" donation types appearing in 2019 and "pledges" and "recurring" donations emerging in 2020. However, recurring donations have consistently remained less than 0.5% of all individual contributions.
- The total count of donors has declined from 2020–2024, though 2022 and 2023 saw temporary increases, likely influenced by post-COVID recovery, strategic campaigns, and rebranding efforts.

2. Membership and Organizational Changes (2012–2024)



- Membership (as measured by first-time donors) showed a sharp drop in 2019 following rebranding, and remained at historically low levels through 2023–2024..
- Significant organizational policy changes (e.g., inclusion of girls in 2018, rebranding in 2019, expanded leadership eligibility) correspond to noticeable shifts in both membership and donation activity, especially the sharp decline in 2019.
- COVID-19 (2020–2022) and bankruptcy proceedings (2020–2023) had substantial financial and membership impacts. Donation totals fell sharply in 2019 and only partially recovered post-pandemic, while new member enrollment—especially in Arizona—remained low despite national efforts to restore trust and visibility.

3. **Synthetic Data**
    - **DonorProfile.csv:** Synthetic dataset representing individual donor profiles, including demographics and giving behavior. Useful for segmentation and personal development in fundraising strategy. The synthetic dataset was constructed with 100,000 donor records, each enriched with realistic behavioral and demographic attributes, including:
        - **Donation Frequency** (0–20)
        - **Average Historical Donation**
        - **Digital Engagement Score** (0–100)
        - **Last Donation Recency** (in months)
        - **Wealth Index** (0–100)
        - **Event Participation Level** (encoded ordinally)
        - **Custom Segment Label** (axio_segment_number, 0–14)
    - **ML-Powered_Propensity_to_Donate_Prediction_for_Grand_Canyon_Council.csv**: Contains machine learning-generated propensity scores predicting the likelihood of future donations. Based on behavioral and historical donor features.
    - **normalised_dataframe_segments.csv**: Normalized and segmented synthetic data prepared for machine learning models. Used for donor clustering and behavior analysis.
    - **Donation - AI Use Cases for Boy Scouts of America_Feb 2025.xlsx:** An Excel workbook outlining AI use cases in fundraising, including synthetic data examples and model outputs. Supports exploration of donation prediction, segmentation, and engagement strategies.

By incorporating these additional data sources, we can strengthen our analysis and develop more effective donor engagement strategies.

4. The Tools

By combining conceptual tools like RFM analysis and behavioral economics with technical tools such as SQL, Python, K-means clustering, predictive modeling, and Tableau, this project will leverage proven methods and technologies to optimize donor engagement strategies. By building on prior successes and using data-driven insights, the Grand Canyon Council will be able to refine its approach to donor management and drive sustainable growth in donations.

Conceptual Tools:

1. RFM Analysis (Recency, Frequency, Monetary Analysis)

Literature Review: RFM analysis is a well-established method in donor segmentation, which has been widely applied in the nonprofit sector to effectively understand donor behavior. Sharp (2010) highlights that RFM segmentation helps organizations identify key donor groups based on their past donation behavior, enabling tailored outreach strategies. For example, frequent and high-value donors are often prioritized for exclusive engagement opportunities, while low-value or infrequent donors may require different engagement strategies to boost their contributions.

Case Study: A great example of RFM analysis in practice comes from Cal Farley's Boys Ranch, a nonprofit focused on at-risk youth, which successfully increased its ROI by 192% using this methodology. This case demonstrates how understanding donor behavior through RFM segmentation can significantly enhance fundraising effectiveness. Although the KPI used in this example was ROI, which differs from the key metrics GCC primarily focuses on—growth of donors and overall fundraising, discussions with GCC leadership have indicated a strong interest in improving ROI as well. Given the similar nonprofit nature of both organizations, applying RFM analysis to GCC's donor data presents a promising opportunity to optimize donor engagement and increase fundraising efficiency.

Application in the Project: In this project with the Grand Canyon Council, RFM analysis will serve as the foundation for understanding donor behavior by segmenting donors into three categories:

- Recency (R): Measures how recently a donor has contributed. Donors who have donated recently are more likely to respond positively to future engagement. This metric will help identify active donors who should be targeted for follow-up.
- Frequency (F): Measures how often a donor contributes. More frequent donors demonstrate higher engagement, and they are vital for sustaining revenue. These donors will be prioritized for engagement strategies aimed at increasing their lifetime value.
- Monetary (M): Measures the monetary value of donations. High-value donors are critical to fundraising success and will be given special attention through personalized strategies to encourage larger contributions or more frequent donations.

By applying RFM analysis, the Grand Canyon Council can enhance their segmentation strategy, ensuring targeted communication to various donor groups based on their past donation behavior.

2. Behavioral Economics

Literature Review: Behavioral economics emphasizes that human decisions are often influenced by psychological, social, and emotional factors, rather than pure rationality (Thaler & Sunstein, 2008). In the context of charitable giving, donors are not always motivated by rational cost-benefit analyses but are influenced by factors like social norms, reciprocity, and emotional fulfillment. For instance, donors may feel a sense of belonging or achievement by contributing to a cause, which increases their likelihood of recurring donations.

Case Study: Researchers at the University of Chicago Booth School of Business conducted a large-scale, two-year natural experiment involving hundreds of thousands of participants to examine how behavioral economics principles can be leveraged to boost charitable giving. The study found that donor behavior is influenced by various psychological factors, such as social recognition, reciprocity, and emotional fulfillment. By carefully designing messages and prompts, nonprofits were able to effectively influence donor decisions, leading to increased donation amounts and frequency. This study highlights the importance of understanding donor psychology and behavior to optimize fundraising strategies. This case supports the application of behavioral economics in the Grand Canyon Council project, reinforcing the value of engaging donors through intrinsic motivations and personalized communication strategies.

Application in the Project: In the Grand Canyon Council project, behavioral economics will be used to inform the design of engagement strategies that appeal to donors' intrinsic motivations. For example, by understanding that donors may be driven by social recognition or self-esteem needs (aligned with Maslow's Hierarchy of Needs), the nonprofit can craft personalized messages that tap into these emotions, making donors feel valued and understood. Such strategies may encourage more substantial and consistent contributions, especially when donors feel a deeper connection to the cause.

Technical Tools:

1. SQL for Data Extraction, Transformation, and Cleaning

- Data Extraction: Queried donor records from the gift transaction table and constituent table using SQL (e.g., JOIN, GROUP BY, WHERE).
- Data Transformation & Cleaning: Removed duplicates, handled missing values, and normalized fields like donation amount, frequency, and recency.

- Feature Engineering: Aggregated data at the donor level to compute total donations, donation frequency, and most recent donation year for use in RFM analysis.

2. Python for Data Analysis

- **RFM Score Computation:**
    1. Used pandas to calculate Recency, Frequency, and Monetary metrics from donation data (2019–2023).
    2. Standardized and ranked RFM components into 1–5 quintiles, combining them into composite RFM segments (e.g., 555 for VIP donors).
- **RFM Segment Aggregation:**
  Created both string-based RFM segments (e.g., 555 for VIP donors) and total score (e.g., 15 for VIP donors) for use in high-precision targeting and modeling.

3. Clustering Algorithms (K-means)

- Applied KMeans from scikit-learn on RFM features to uncover donor behavioral patterns:
  (e.g., High-Value Loyalists: Frequent and high-amount givers, New High Potential: Recent donors with large or moderate donations, At-Risk Donors: Previously active but now inactive donors.)
- Determined optimal cluster number using elbow method and silhouette score, then profiled each cluster to inform strategy.

4. Predictive Modeling (Random Forest)

- **Regression Model:**
  Predicted donation amount for 2025 using RandomForestRegressor trained on 2023 behavioral and RFM features.
- **Model Evaluation:**
    1. Used R^2, RMSE, and feature importance to evaluate and explain model performance.
    2. Created scenario forecasts assuming different retention rates among top predicted donors
       .

5. Data Visualization and Dashboards

- Python (Matplotlib & Seaborn): Used to create line charts, bar plots, scatter plots, and boxplots for RFM distribution, donation trends, and model results.
- SQL + Python: Extracted and cleaned donor data with SQL, then visualized trends and features in Python.
- Streamlit: Built an interactive dashboard to upload data, predict segments and donations, and display insights with dynamic charts.

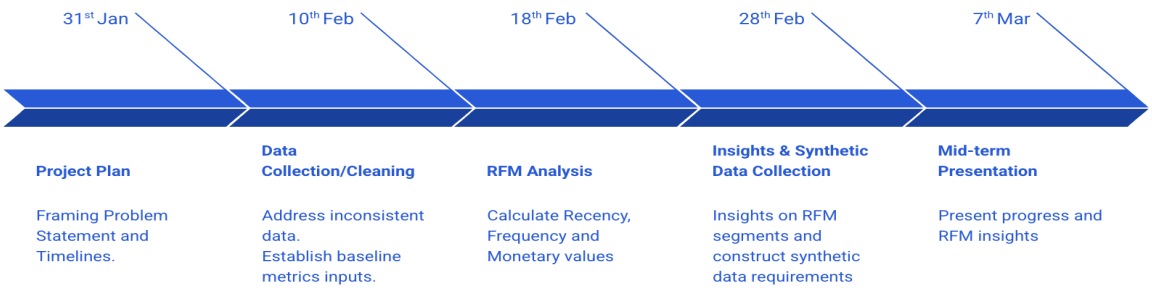
5. Execution

Index	Task	Owner (team member)	Start Date	End Date	Dependency (Task Indices)	Status
1. Project Planning						
1.1	Problem statement	Austin, Darshan	1/15	1/31	-	Completed
1.2	Project Timeline development	Zih Han, Julis	1/15	1/31	1.1	Completed
2. Data Collection and Cleaning						
2.1	Establish baseline metrics for RFM inputs	Zih Han	1/31	2/3	-	Completed
2.2	Address inconsistent data formats	Julis, Zih Han	2/3	2/5	2.1	Completed
2.3	Remove duplicate records	Zih Han	2/5	2/7	2.1	Completed
2.4	Address missing RFM-critical fields	All	2/7	2/10	2.2, 2.3	Completed
3. RFM Analysis						
3.1	Calculate Recency, Frequency, and Monetary values	Austin, Zih Han	2/11	2/14	2.4	Completed
3.2	Segment donors using RFM scores	Zih Han	2/15	2/18	3.1	Completed
4. Insights and Strategy Development						
4.1	Develop tailored engagement strategies for RFM segments	Julis, Zih Han	2/18	2/28	3.2	Completed
5. Mid-term Presentation						
5.1	Present progress and RFM insights	All	3/7		1.1 - 4.1	Completed
Spring Break (3/8 - 3/15)						
6. Donor Segmentation using Clustering						
6.1	Identify key features and scaling the data for fair clustering	Zih Han	3/16	3/20	2.1 - 2.4	Completed
6.2	Applying K-Means Algorithm to the data	Zih Han, Darshan	3/21	3/25	6.1	Completed
6.3	Interpret and Visualize the Clusters	Zih Han	3/25	3/28	6.3	Completed

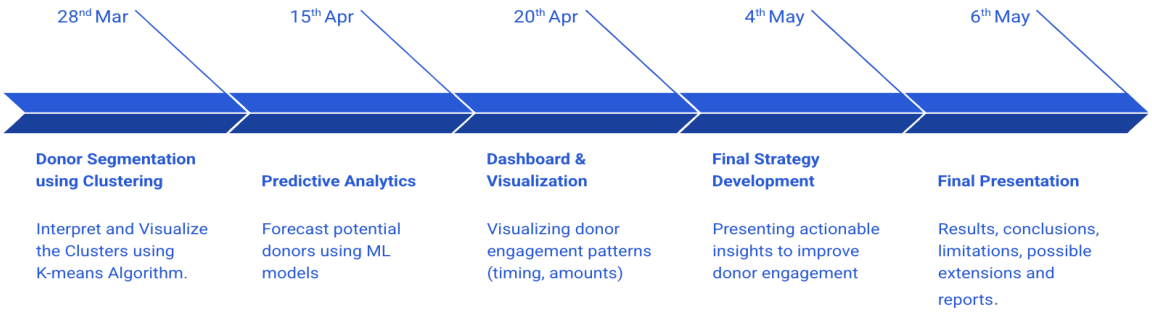
7. Predictive Analytics for Donor Behavior						
7.1	Implement possible Decision Tree/Logistic Regression model to understand Donor retention	Zih Han	3/29	4/8	2.1 - 2.4	Completed
7.2	Forecast potential donors who will donate next year	Zih Han	4/9	4/15	7.1	Completed
8. Dashboard & Visualization						
8.1	Visualizing donor engagement patterns (timing, amounts)	Zih Han Darshan	4/12	4/14	2.1	Completed
8.2	Poster creation (visual integration, storytelling, design layout)	Zih Han Austin	4/7	4/14	1.1 - 8.1	Completed
9. Final Strategy Development						
9.1	Refine strategies for new, lapsed, and repeat donors	Julis	4/21	4/23	6.1 - 8.2	Completed
9.2	Presenting actionable insights to improve donor engagement	Zih Han	4/24	4/26	6.1 - 9.1	Completed
9.3	Develop final recommendations to achieve 2.5% donation growth.	All	4/27	5/4	6.1 - 9.2	Completed
10. Final Presentation						
10.1	Presenting problem statement, business value, scope, assumptions, approach, results and conclusions, limitations, possible extensions	All	5/6		1.1 - 9.3	Completed
10.2	Final Report	Zih Han Darshan	5/6		1.1 - 9.3	Completed

All relevant project materials are organized and accessible in the shared project folder linked below:
[Grand Canyon Council Capstone Project]

First Quarter



Second Quarter



Project stakeholder meeting:

Index	Meeting Objective	Meeting Attendees	Start DateTime	End DateTime	Next Steps Identified (possibly a task index in Project Execution Plan)
1	Kick-off meeting: Introduce project goals, and responsibilities. Align on the overall objectives and clarify any initial questions.	All Project Team Members, Client, Advisor	12/30, 8:00 AM	12/30 9:00 AM	Task 1 - Project Planning
2	Discuss data collection and cleaning progress, project plan, identify potential data quality issues, and finalize data requirements.	All Project Team Members, Client, Advisor	1/28 Internal Meeting	1/28 Internal Meeting	Task 2.1 – 2.4: Finalize Data Cleaning Strategy
3	Mid-term check-in: Review the problem statement, business value, intermediate results, and next steps ahead of the Mid-term Presentation.	All Project Team Members, Client, Advisor	3/4 12:00 PM	3/4 12:45 PM	Task 5: Preparation for Mid-term Presentation
4	Review of data analysis, segmentation, and model building; address any blockers or challenges. Discuss data-driven insights for engagement strategies.	All Project Team Members, Client, Advisor	3/18 Internal Meeting	3/18 Internal Meeting	Task 3, 4: Strategy Development
5	Final presentation prep: Review findings, results, recommendations, and limitations before the final presentation.	All Project Team Members, Client, Advisor	4/30 9:00 AM	4/30 9:30 AM	Task 10: Preparation for Final Presentation

6. Risks and Issues

Risks and challenges may impact the project's success. Below are the key risks, their potential impact, and how they can be managed effectively.

1. **Incomplete or Inaccurate Donor Data**
   - **Risk:** Some donor records have missing or inconsistent data, such as missing donation amounts, engagement history, or outdated contact details.
   - **Impact:** Poor data quality can lead to inaccurate insights, making it difficult to personalize donor outreach or predict future donations.
   - **Mitigation:**
     - Perform thorough data cleaning before analysis.
     - Identify and fill missing values using statistical methods or outreach to the client for clarification.

2. **Variability in Donor Behavior**
   - **Risk:** Donor contributions are influenced by economic conditions, personal priorities, and external events (e.g., inflation, recessions, or changes in donor interests).
   - **Impact:** If engagement strategies do not account for these changes, predictions about donor behavior may be inaccurate.
   - **Mitigation:**
     - Use historical trends to identify external factors affecting donations.
     - Implement real-time tracking of donor behavior to adjust strategies dynamically.

3. **Predictive Model Accuracy and Usability**
   - **Risk:** Predictive models used to forecast donor behavior may be inaccurate or overfitting if trained on biased or incomplete data.
   - **Impact:** If the model's predictions are unreliable, fundraising teams may invest in the wrong engagement strategies.
   - **Mitigation:**
     - Validate the model using historical data before deployment.
     - Use multiple performance metrics (e.g., accuracy, precision, recall) to assess reliability.

4. **Differences Between Individual and Corporate Donors**
   - **Risk:** Individual and corporate donors have different motivations for giving. A one-size-fits-all approach may reduce effectiveness.
   - **Impact:** Poor segmentation may fail to retain individual donors or engage high-value corporate sponsors.
   - **Mitigation:**
     - Develop tailored engagement strategies for each donor type.
     - Use segmentation analysis to identify patterns in donor behavior.

5. **Ethical responsibilities accountability to the Society**
    - **Positive Impact:** Increases donor contributions, ensuring more youth in Arizona benefit from Scouting programs that foster leadership and community engagement.
    - **Concern:** Favoring wealthier donors could create an equity gap in engagement efforts.
    - **Mitigation**: Ensure inclusive donor outreach regardless of contribution size.

6. **Ethical responsibilitiesAccountability to the Organization**
    - **Positive Impact:** Enhances fundraising efficiency, enabling better resource allocation and data-driven decision-making.
    - **Concern:** Misuse or leakage of donor data could harm organizational trust and reputation.
    - **Mitigation:** Enforce strict data privacy protocols to protect sensitive donor information.

7. Data Challenges and Solutions

During the analysis of donor and constituent datasets, several data quality and completeness issues were identified. To address these limitations and ensure robust analytical outcomes, synthetic data techniques were applied. The following summarizes the key challenges encountered and the solutions implemented:

1. **Sparse Donation and Engagement Data**
   Many donor records lacked sufficient transaction history or consistent engagement activity. This limited the ability to analyze trends or build predictive models based on real behavior.
   Solution: Synthetic donor profiles were generated to simulate realistic donation and engagement patterns. These synthetic records filled critical data gaps, enabling comprehensive trend analysis and enhancing the training dataset for modeling purposes.

2. **Inconsistent Constituent Information**
   The constituent dataset exhibited considerable variation in completeness, particularly in key demographic and engagement-related fields. These inconsistencies posed challenges for donor segmentation and personalized outreach.
   Solution: Synthetic attributes such as estimated wealth indicators and engagement levels were programmatically added. This enrichment process created a more uniform and analytically useful dataset, supporting more accurate segmentation and targeting.

3. **Limited Behavioral Depth**
   The available data captured only basic interaction details, lacking in-depth behavioral insights such as giving motivation, frequency patterns, or lifecycle stages. This limited the development of nuanced models and strategies.
   Solution: Synthetic data augmentation techniques were used to introduce additional behavioral dimensions. This enabled more advanced modeling approaches and allowed for precise donor segmentation and the development of targeted engagement strategies.

8. RFM

RFM (Recency, Frequency, Monetary) analysis was used to segment donors based on their giving behavior. This method provides a powerful framework for understanding donor engagement, value, and

potential. The following subsections outline how scoring was applied and how behavior-based segments were derived from the scores.

1. **RFM Scoring Methodology**

   The RFM model assigns each donor a score from 1 to 5 in three dimensions:

   - **Recency**: How recently a donor made a gift.
   - **Frequency**: How often a donor contributed within 4 years.
   - **Monetary**: The total amount a donor contributed during 4 years.

Donors are ranked into quintiles for each metric annually. The most favorable 20% receive a score of 5, while the least favorable 20% receive a score of 1.

Metric	Score 5	Score 4	Score 3	Score 2	Score 1
Recency	Most recent 20%	20%-40%	40%-60%	60%-80%	Oldest 20%
Frequency	Most frequent 20%	20%-40%	40%-60%	60%-80%	Least frequent 20%
Monetary	Top 20% spenders	20%-40%	40%-60%	60%-80%	Bottom 20% spenders

For example:

- A **recency score of 5** means the donor gave within the most recent 20% of the year.
- A **frequency score of 5** places the donor among the top 20% in giving frequency.
- A **monetary score of 5** represents the top 20% of donors in terms of contribution value.

RFM scores are combined into a three-digit code (e.g., 555, 114), forming the basis for segmentation.

2. **Behavioral Segmentation Based on RFM**

Using the RFM scores, donors were classified into distinct behavioral segments, each reflecting a unique combination of engagement patterns and donation value

Segment Definitions:

Segment	RFM Score Rule	Behavior Description
VIP Donor	555	Top scores in all three metrics; recently active, frequent, and high-value donors

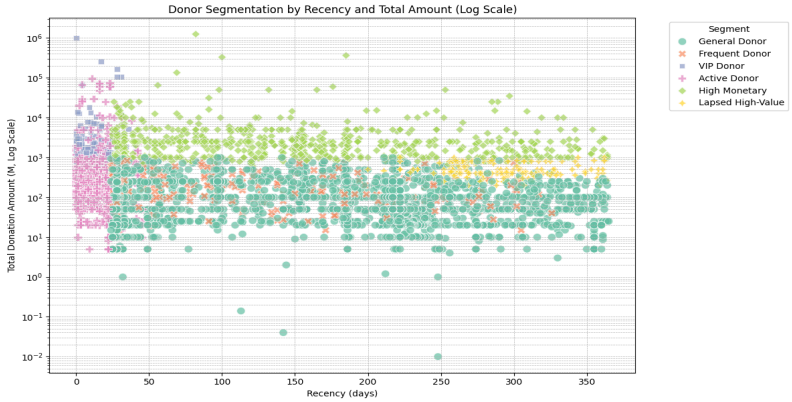
High Monetary Donor	Monetary = 5 (e.g., 115)	Consistently high donation amounts, but not necessarily recent or frequent
Frequent Donor	Frequency = 5 (e.g., 151)	Regular engagement, typically with moderate donation value
Active Donor	Recency = 5 (e.g., 511)	Recently engaged but may not be frequent or high-value
Lapsed High-Value	114	Historically high-value donors with recent inactivity
High Recency, Low Value	531	Recently gave, but with low value and frequency
General Donor	None of the above	Largest group; generally lower in all three metrics, potential for nurturing

3. Behavioral Insights from Segmentation

A visual analysis using a scatter plot of **Recency (days)** vs. **Total Donation Amount (log scale)** revealed key behavioral traits:

- **VIP Donors** cluster in the upper-left, representing high-value, recently engaged supporters.
- **High Monetary Donors** rank high on the donation scale but may appear further right, showing inconsistent recency.
- **Frequent Donors** are spread horizontally with moderate donation amounts but high consistency.
- **General Donors** form the majority, concentrated in low-value, low-recency areas—ideal for growth and engagement.
- **Lapsed High-Value Donors** stand out with high historical value but older donation recency

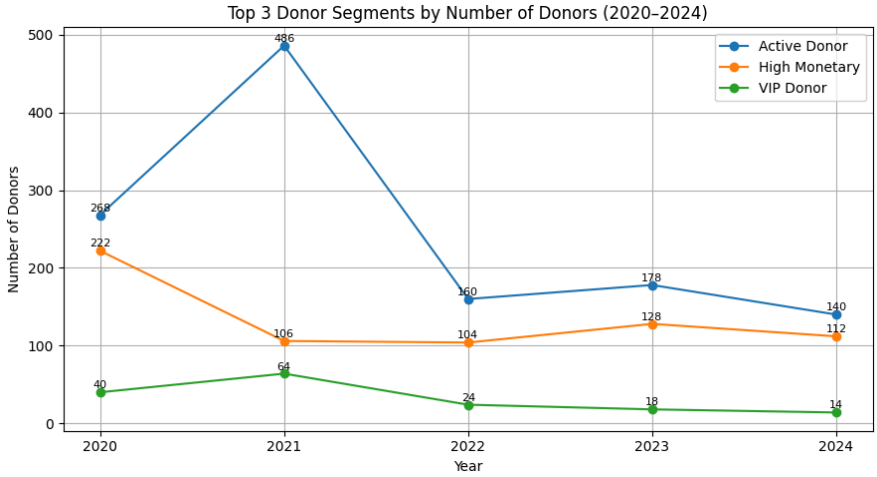
This segmentation allows targeted engagement strategies: retain VIPs, reactivate lapsed high-value donors, and nurture general donors toward higher value or frequency.



9. Preliminary Analysis

This section presents a deeper exploration of donor trends observed between 2020 and 2024. It expands on the previously established problem by analyzing changes in donor behavior and highlights emerging opportunities for strategic fundraising.

1. Problem Deep Dive



An in-depth analysis of donor segments reveals significant declines in both participation and contribution levels across key donor categories:

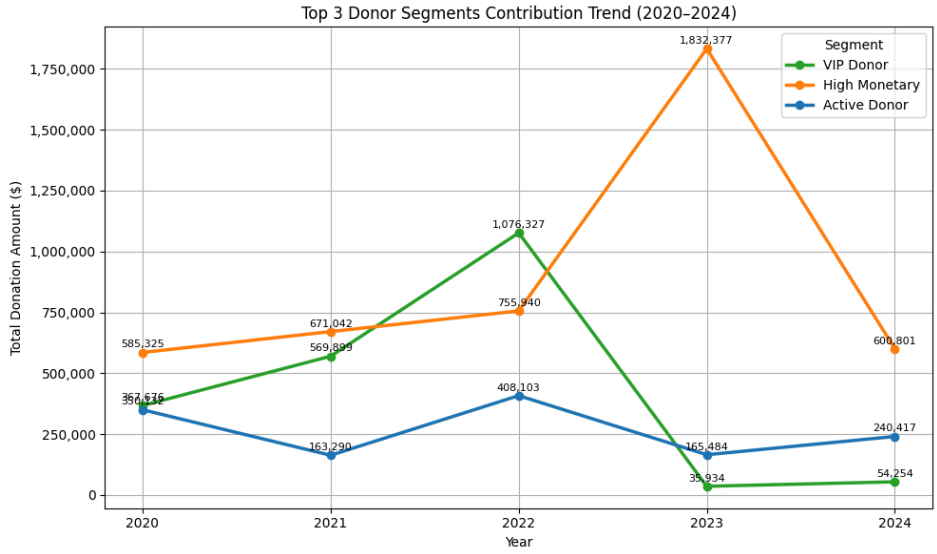
- **Active Donors** have decreased by more than **70%** since their peak in 2021, indicating a major drop in general engagement.
- **High Monetary Donors** — individuals contributing significantly higher donation amounts — have declined by nearly **50%** compared to 2020 levels. While there was a temporary recovery in 2023, the trend remains downward.
- **VIP Donors**, representing the highest-value and most engaged individuals, have experienced the steepest attrition, dropping by over **75%** from their 2021 peak.

These trends, visualized in the donor segment chart (Figure X), suggest systemic issues in donor retention and engagement strategies, particularly among high-value segments. The implications are especially severe for fundraising sustainability, as the loss of top-tier donors results in disproportionately large revenue impacts.

Key Challenges Identified:

- Donor base shrinkage, particularly within high-value segments critical to fundraising efforts.
- Severe VIP donor attrition, risking long-term relationship erosion.
- Sparse and inconsistent data, which undermines the ability to predict donor behavior or proactively manage attrition.

2. **Opportunity Identification**



Despite the evident decline, donor contribution trends highlight a strategic opportunity to recover and grow through focused engagement of high-value donors:

- In **2023**, **High Monetary Donors** contributed over **90%** of the total donation volume. This indicates that, although their numbers have declined, their financial impact remains significant.
- **VIP Donors**, though fewer in number by 2024, accounted for over **50%** of total donations in **2022**, highlighting their potential when engaged effectively.
- These segments continue to demonstrate high return on engagement, making them ideal targets for strategic re-engagement and retention efforts.

Strategic Focus Areas:

- **Prioritize re-engagement** of High Monetary and VIP Donors through personalized communication, stewardship, and targeted campaigns.
- **Implement predictive modeling** to identify lapsed or at-risk donors early, enabling preemptive outreach.
- **Leverage enriched and synthetic data** to supplement gaps in behavioral insight, allowing for more accurate segmentation and strategy development.

By shifting resources toward sustaining and growing relationships with these high-value segments, the organization can drive consistent revenue growth and improve overall donor lifetime value

10. K-Means Cluster

To validate the donor segmentation derived from RFM scoring, a **K-means clustering analysis** was conducted using unsupervised machine learning techniques. This approach independently grouped donors based on behavioral features such as **recency** and **donation amount**.

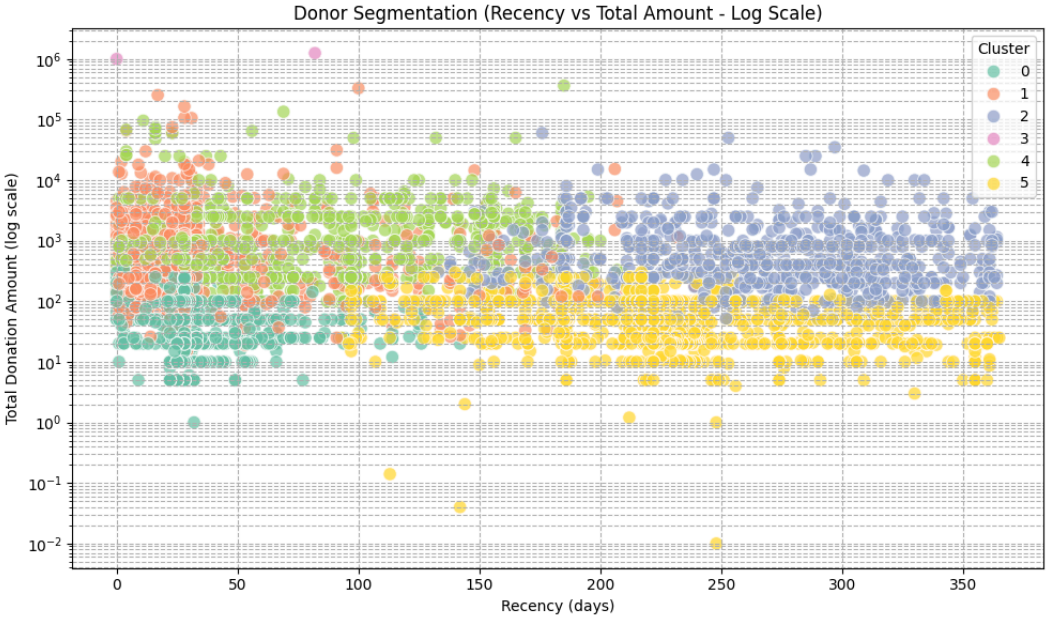
Objective

The primary goal of applying K-means clustering (with **k=6**) was to:

- Validate the RFM-based segmentation structure using an unsupervised method.
- Uncover any additional natural groupings not previously identified.
- Strengthen confidence in segmentation strategies through data-driven alignment.

Findings

- **Alignment with RFM Segmentation**: The K-means model naturally identified six donor clusters that closely mirrored the behavioral patterns discovered in RFM segmentation. This alignment confirms the robustness of the initial segmentation logic.
- **Distinct Donor Profiles Emerged**: Each cluster represented a unique combination of **recency** and **donation value**, reinforcing the diversity in donor engagement styles.
- **Supports Strategic Formulation**: The clustering results provide further support for targeted strategies tailored to distinct donor types, from re-engaging lapsed high-value donors to cultivating frequent but lower-value contributors.



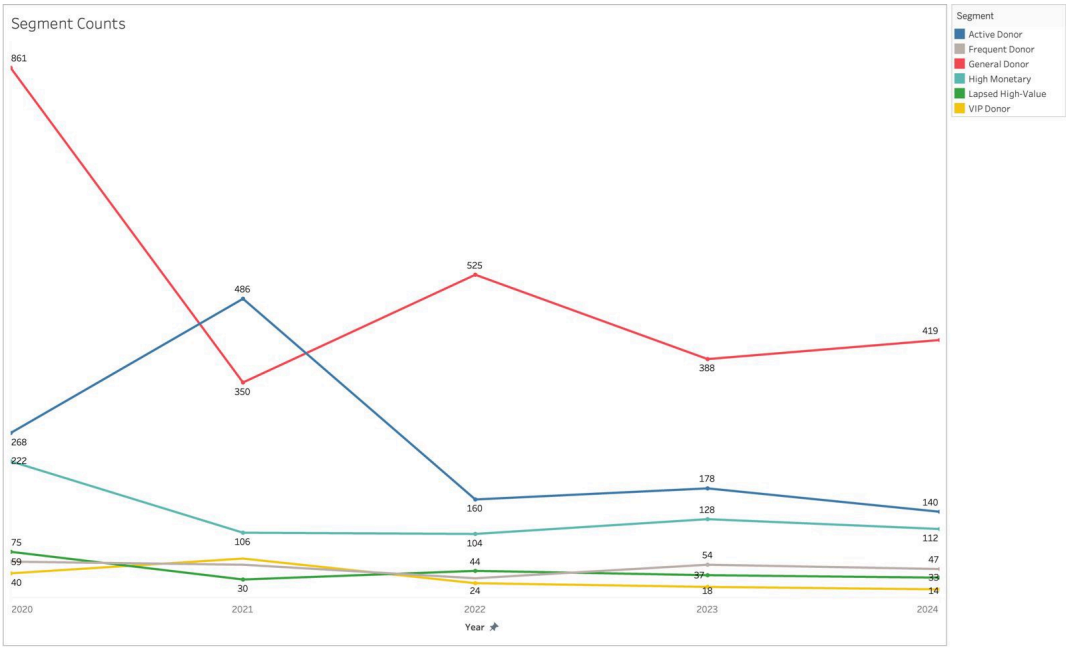
This exploration using unsupervised learning not only validates the original segmentation but also enhances the reliability of data-driven decision-making for donor engagement strategies.

11. Donor Trends Over Time

An examination of donor segment counts from 2020 to 2024 reveals substantial instability across various donor types, highlighting the need for proactive engagement and retention strategies.

Donor Instability Signals Urgency for Action

- **General Donors** experienced a dramatic drop from **861 in 2020** to **350 in 2021**, reflecting early instability in the donor base. Although there was partial recovery (525 in 2022), counts have not returned to initial levels, ending at **419 in 2024**.
- **Active Donors** peaked at **486 in 2021** but declined steadily in the following years, falling to **140 by 2024**. This trend indicates a concerning decrease in recent engagement, potentially tied to donor fatigue or lack of follow-up.
- **High Monetary Donors** declined sharply from **222 in 2020** to **106 in 2021**, and have since stabilized in the **100–130 range**. Although still valuable, this stagnation signals limited growth in high-value donor acquisition.
- **Frequent Donors** and **Lapsed High-Value Donors** are showing **gradual upward trends**, increasing from 2021 to 2024. This suggests potential for reactivation and deeper engagement, especially among previously high-value contributors.
- **VIP Donors** consistently remain a **very small group**, with no significant growth over the years. This segment may be difficult to scale without highly tailored acquisition and stewardship strategies.



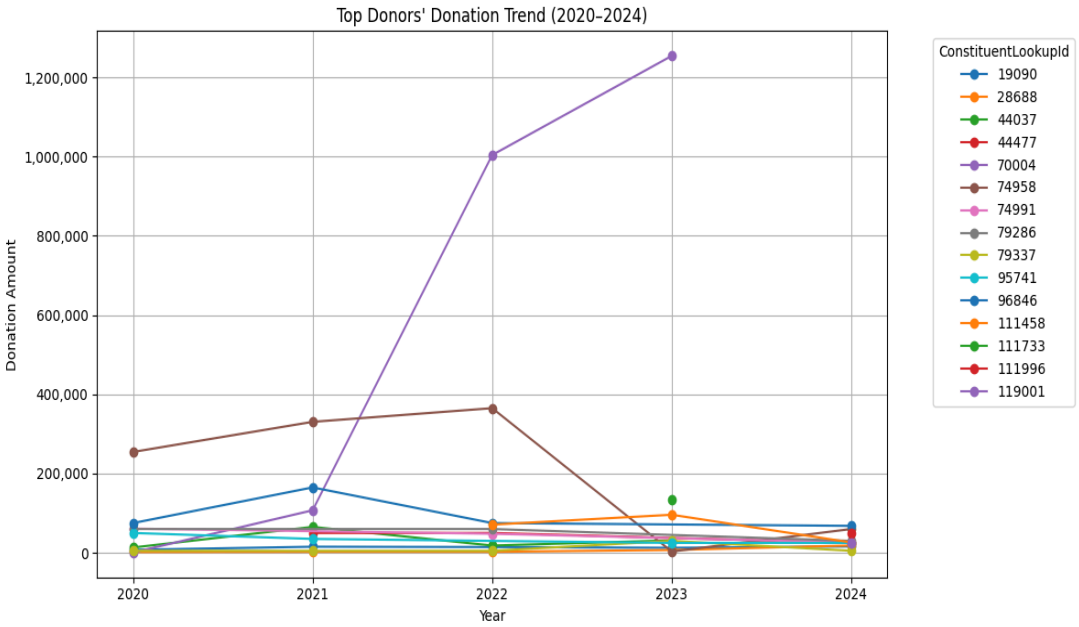
Strategic Implications

- The overall donor landscape is marked by **fluctuations and contractions**, especially in engagement (recency) and high-value contribution.
- Growing segments like **Frequent** and **Lapsed High-Value** donors present **re-engagement opportunities**.
- The stagnation in VIP and High Monetary groups indicates a need to **revamp high-tier donor acquisition** and retention programs.
- The sharp decline in General and Active donors underscores the importance of **maintaining the donor funnel**, from early engagement to long-term stewardship.

12. Strategic Profiling of High-Value Donors

To enhance donor management and maximize retention, high-value donors were profiled based on their historical donation behavior from 2020 to 2024. This behavioral profiling enabled the identification of distinct engagement patterns and informed targeted strategies for each segment.

Donor Segmentation Overview



High-value donors were segmented into four key categories based on their annual giving trends:

Category	Behavior Pattern	Sample Donor IDs
Consistent Growers	Donated every year with steady or increasing contribution amounts	28688
Returning Donors	Skipped one or more years but resumed giving in recent periods	44477, 74991792, 86, 95741, 96846, 11199611, 9001
Declining Supporters	Donated across multiple years, but with a notable drop in recent value	79337, 111458
Lapsed Donors	Were active donors but made no contribution in 2024	44037, 70004, 111733

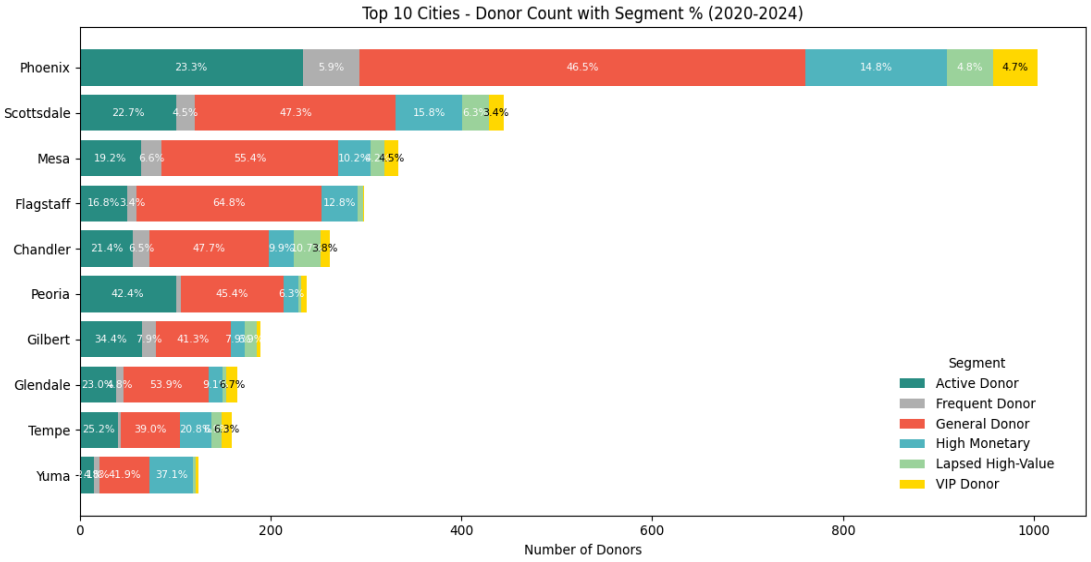
A multi-year donation trend chart illustrates how these segments differ in terms of contribution trajectory, highlighting both potential risk areas and opportunities

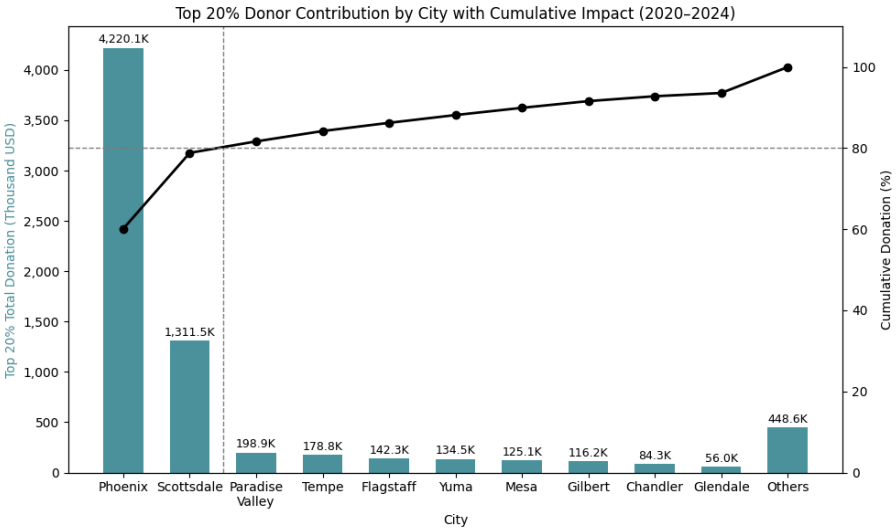
Tailored Strategies by Donor Type

- **Consistent Growers**
    - **Profile:** Reliable donors with steady or increasing annual contributions.
    - **Strategy:** Prioritize with long-term stewardship—offer recognition, naming opportunities, and inclusion in VIP programs.
- **Returning Donors**
    - **Profile:** Previously lapsed donors who resumed giving.
    - **Strategy:** Reinforce re-engagement with personalized outreach and gratitude; understand triggers that brought them back.
- **Declining Supporters**
    - **Profile:** Donors with noticeable decline in recent donation amounts.
    - **Strategy:** Send re-engagement messages and brief surveys to explore disengagement; offer flexible giving options.
- **Lapsed Donors**
    - **Profile:** No donations made in 2024 despite past contributions.
    - **Strategy:** Launch reactivation campaigns using impact-driven storytelling, incentives, and contact update prompts.

13. Geography-Driven Strategy for Donor Engagement

Analyzing donor distribution by city revealed that a large share of high-value contributions comes from a small set of geographic areas. This presents a significant opportunity to localize fundraising strategies for higher efficiency and impact.





Strategic Focus on High-Contributing Cities

- **Geographic Concentration**: Nearly **80% of total high-value donations** between 2020 and 2024 originated from **Phoenix** and **Scottsdale**, demonstrating a strong donor presence and philanthropic capacity in these two cities.
- **Targeted Growth Potential – Phoenix**:
    - Phoenix houses the **largest concentration of General Donors**, which presents a substantial opportunity for **donor upgrades** through re-engagement campaigns, stewardship initiatives, and value-tier migration.
    - The city also exhibits diversity in donor types, suggesting scalability of outreach strategies.
- **VIP Retention Priority – Scottsdale**:
    - Scottsdale has a **notably high density of VIP and High Monetary donors**, making it an essential focus for **retention and loyalty programs**.
    - With a higher median income, education level, and household value compared to other cities, Scottsdale's donor base is primed for premium engagement strategies (e.g., leadership events, recognition tiers).

Supporting Socioeconomic Insights: The demographic and socioeconomic profile of Phoenix, Scottsdale, and Mesa further reinforces strategic targeting:

City	Median Income	Bachelor's Degree or Higher	Median Home Value	Poverty Rate
Scottsdale	$107,372	61.5%	$709,900	7.1%
Phoenix	$77,041	32.3%	$381,900	14.3%
Mesa	$78,779	31.5%	$364,300	10.5%

These statistics help tailor messaging and engagement formats that align with community characteristics, e.g., educational-themed appeals in Scottsdale or culturally inclusive campaigns in Phoenix.

14. Predictive Modeling with Historical Data

Objective:
To determine whether historical donation behavior and constituent characteristics could effectively predict future giving, and thus guide strategic decision-making.

Approach:

- Used donation data from **2020–2023**, combined with **2023 RFM scores** and **constituent data**.
- Trained a **Random Forest regression model** to predict individual donation amounts for 2024.

Results:

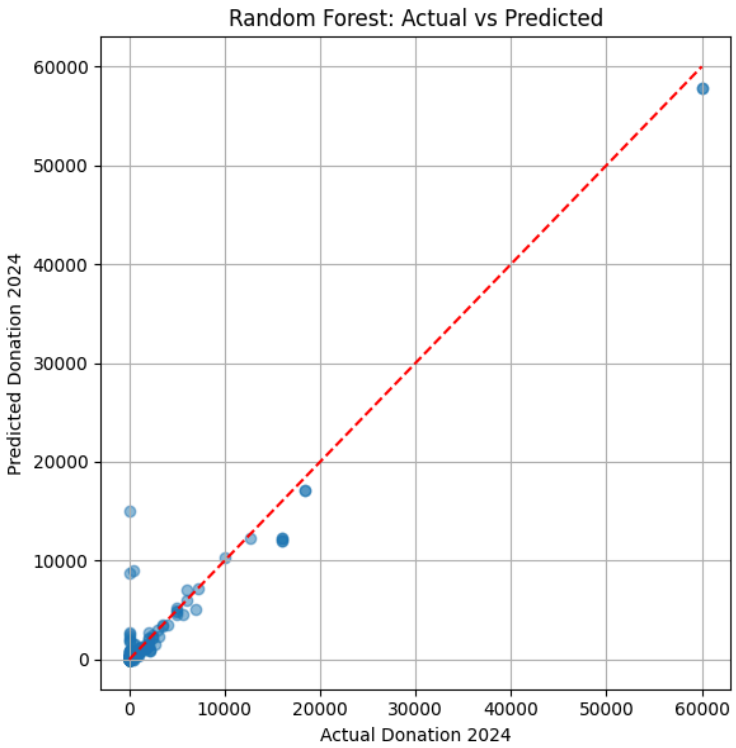
- The model demonstrated **strong predictive performance**:
    - **Training R^2**: 0.9803 | **Testing R^2**: 0.9454

- **Training RMSE**: 481.64 | **Testing RMSE**: 1105.44

- The scatter plot of actual vs. predicted donation amounts shows **tight alignment** with minimal variance from the ideal fit line.

Implications:

- Validated that future giving can be well explained by **historical donor behavior by 94.5%**.
- Reinforces the value of behavioral segmentation and constituent analytics in driving targeted fundraising strategies.



15. Forecast-Driven Donor Strategy for 2025

To optimize fundraising outcomes, predictive modeling was used to identify the top 20% of donors expected to contribute most in 2025. The forecast results helped prioritize donor segments and geographic areas for strategic action. However, model evaluation revealed issues in granularity and personalization, which were addressed through the integration of synthetic data.

1. Forecasting 2025 Donations: Donor & City-Level Insights

Top 20% Predicted Donors

Behavior-based forecasting using 2024 features and RFM scores allowed us to identify the top 20% of donors likely to contribute a disproportionate share of total donations in 2025. These donors represent strong targets for campaign planning.

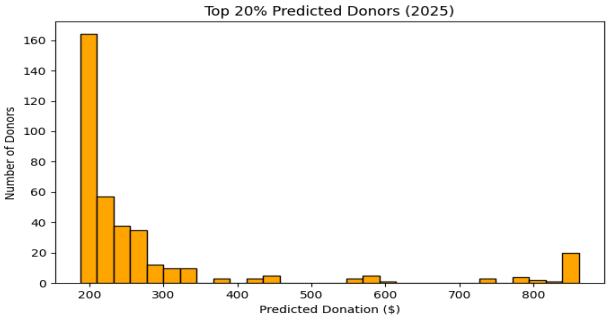
- Most predicted donation values fall below **$1,000**, indicating significant **upgrade potential**.
- This segment is ideal for **personalized stewardship**, upsell strategies, and re-engagement campaigns.

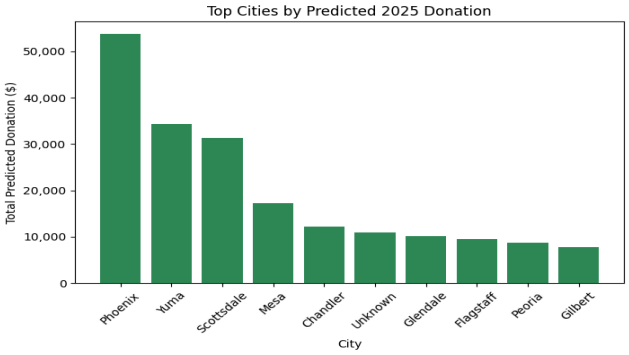
- These donors are projected to contribute a significant share of total donations based on **80/20 principle**, where a minority drives the majority of returns.

Top Cities by Predicted Giving

Geographic trends in predicted donations showed clear regional concentration:

- **Phoenix, Yuma, and Scottsdale** lead in projected giving.
- These cities present optimal **ROI for campaign resources**, staff outreach, and donor events.
- Regional targeting allows for **localized messaging** and efficient allocation of development staff.





2. Evaluation of Prediction Outputs: Identified Weaknesses

While predictive modeling offered strategic foresight, post-modeling analysis revealed critical issues related to prediction quality and personalization.

Prediction Uniformity Among Top Donors

Many top donors were assigned the **same predicted value**, indicating a lack of differentiation:

- Example: Numerous donors from **Yuma, AZ** were assigned **$861.06**.
- This suggests that while the model captured macro-trends, it **failed to personalize predictions at the individual donor level**, limiting targeting effectiveness.

ConstituentLookupId	City	State	predicted_donation_2025
79344	Yuma	AZ	861.9390
79157	Yuma	AZ	861.9390
111835	Yuma	AZ	861.9390
79344	Yuma	AZ	861.0586
79287	Yuma	AZ	861.0586
79344	Yuma	AZ	861.0586
119000	Yuma	AZ	861.0586
79157	Yuma	AZ	861.0586
79287	Yuma	AZ	861.0586
101826	Yuma	AZ	861.0586
118948	Yuma	AZ	861.0586
101826	Yuma	AZ	861.0586
101826	Yuma	AZ	861.0586
20009	Yuma	AZ	861.0586
111835	Yuma	AZ	861.0586
79344	Yuma	AZ	842.9670
20009	Yuma	AZ	842.9670
79287	Yuma	AZ	842.9670
79157	Yuma	AZ	842.9670
20009	Yuma	AZ	842.9670

Repetitive Value Assignments Across the Dataset

Beyond top donors, value duplication was widespread:

- **345 predicted donation values** were shared by multiple donors.
- Example: The value **$3.93** was assigned to **25 different donors**.
- This points to **limited prediction granularity**, undermining the ability to segment donors meaningfully or prioritize outreach.

Predicted Donation ($)	# of Donors Assigned
3.93	25
158.77	24
102.40	20
95.51	19
14.64	19
165.71	18
9.40	17
155.12	17
15.75	16
138.81	16

3. Enhancement Through Synthetic Data Integration

To address these modeling limitations, **synthetic data** was incorporated into the prediction pipeline. This enhanced data quality and expanded the model's ability to simulate diverse donor behaviors

Strategic Value of Synthetic Data

- **Introduces Behavioral Variability**:
  Synthetic donor records simulate a wider range of engagement patterns, addressing the model's tendency to assign repetitive donation values (e.g., $861.06 or $3.93 across many donors). This improved the **uniqueness and personalization** of predictions.
- **Strengthens Donor Segmentation**:
  By introducing slight variations among otherwise similar donor profiles, the model was better able to differentiate between donors. This enhanced **ranking precision** and allowed for more tailored messaging strategies across segments.
- **Supports Scenario Testing and Model Iteration**:
  In cases where real donor behavior data is sparse, outdated, or lacks diversity, synthetic data fills

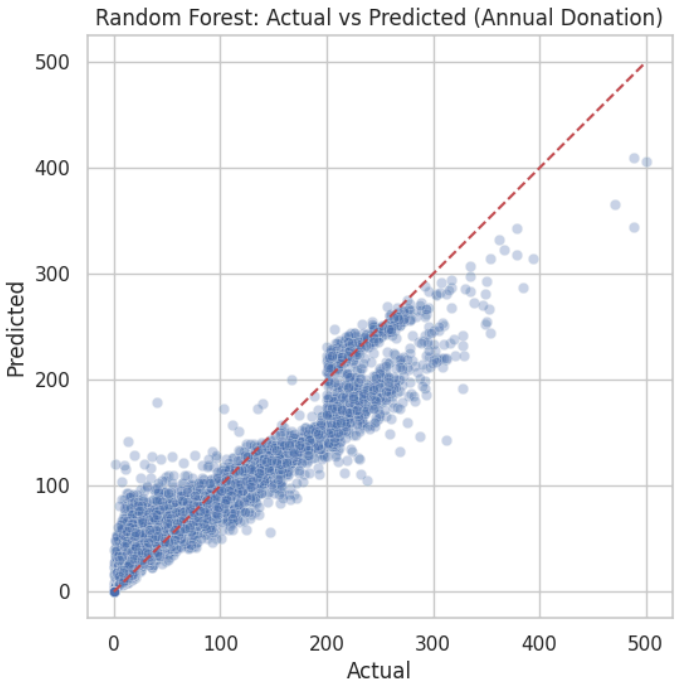
in critical gaps. It enables teams to test and refine models under a variety of hypothetical conditions—especially useful for strategic planning or simulating future campaigns.

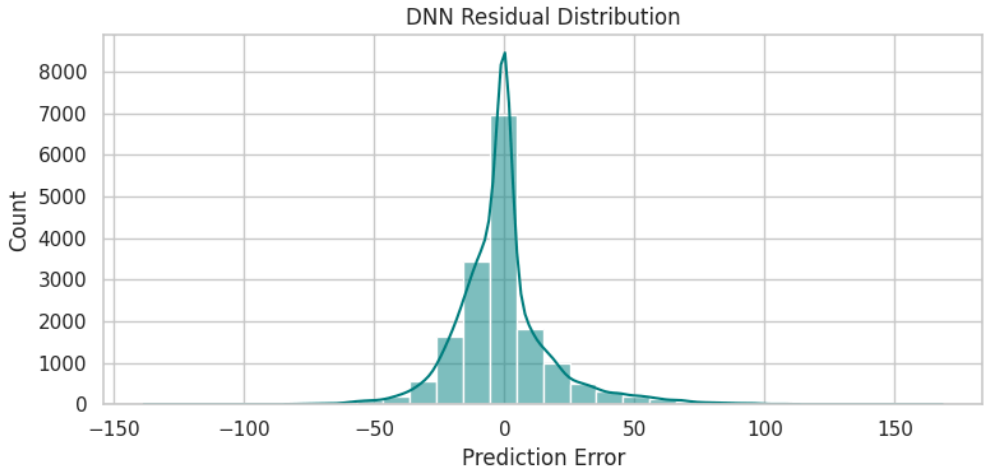
- **Enables Earlier and Safer Forecasting**:
  Because synthetic data can simulate potential future states without relying solely on historical patterns, it allows the organization to **generate forecasts in advance**, de-risking decision-making and enhancing agility in campaign design.

To validate its utility, a regression model was trained solely on the synthetic dataset due to structural differences from real data, making direct integration infeasible. This approach validated the predictive power of synthetic features ($R^2 = 0.9008$), supporting its use for early forecasting and scenario planning when real data is limited.:

  - **Training R^2**: 0.8997 | **Testing R^2**: 0.9008
  - **Training RMSE**: 19.41 | **Testing RMSE**: 19.10

This high performance confirms the dataset's consistency and predictive potential, especially for forward-looking scenarios where real donor data may be limited.

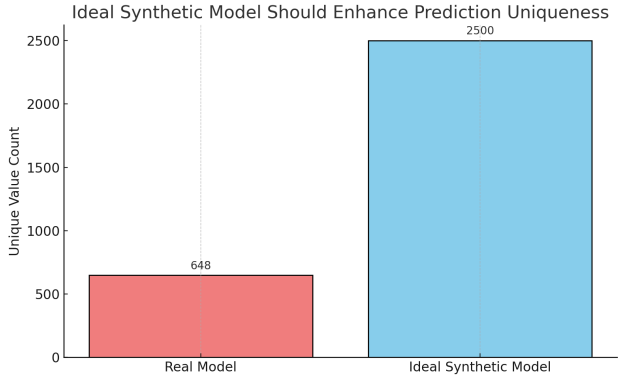




4. Strategic Outcome

Improved Prediction Granularity [Not Yet Achieved]

- The current synthetic model still produces many duplicated predicted values
- Predicted donation range is narrower than the real model (Synthetic: $33–$196 vs. Real: $2.73–$861).
- Donor ranking is flattened, limiting our ability to prioritize within segments.

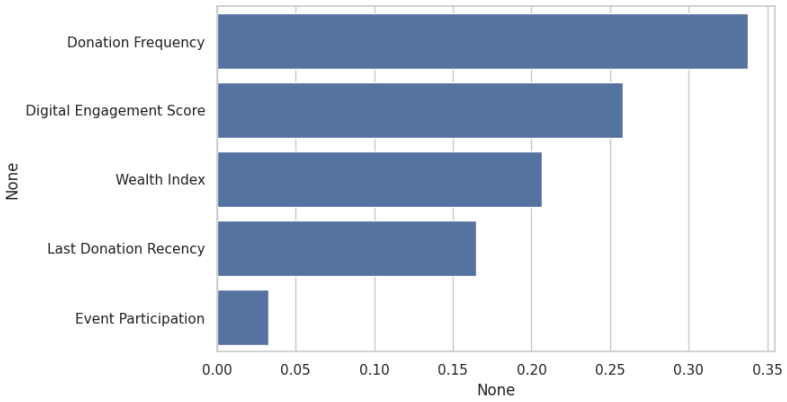


Enhanced Persona-Based Segmentation [Achieved]

- We successfully assigned all donors to 15 marketing-defined personas using RFM + behavioral features.
- Each segment (e.g., Event-Only Non-Donors, Engaged Mid-tier Donors) has a corresponding strategy recommendation.
- This structure allows for audience-specific messaging and campaign planning.

Feature-Driven Insight Extraction [Partially Achieved]

We analyzed feature importance from the synthetic model to identify key behavioral drivers of donation predictions. The top features—**Donation Frequency**, **Digital Engagement Score**, and **Wealth Index**—provided actionable insight and confirmed the model learned meaningful patterns. However, due to the narrow prediction range and repetitive outputs, the impact of other features (e.g., Event Participation) was muted, limiting full interpretability and scenario depth. Richer variation in synthetic profiles will be necessary to unlock the full potential of feature-level decision support.



Forecast-Driven Decision Support [Partially Achieved]

- We used synthetic donor profiles to simulate total donation potential by segment and city.
- This helps in identifying which donor types and regions to prioritize in future outreach.
- Many predicted values are still repetitive, the precision of scenario simulation is limited.
- To simulate realistic donor behavior across regions and time, we need **richer synthetic variation**.
- Currently, the synthetic data lacks **noise and behavioral diversity**, making long-term campaign simulation less reliable.

While our current synthetic model has limitations, particularly in prediction granularity—it offers valuable insights into the potential of synthetic augmentation in nonprofit analytics. Specifically, we achieved strong segmentation and partially supported geographic forecasting, though value-level personalization remains limited due to insufficient variability in the synthetic data.

This work provides a **conceptual foundation** for how synthetic data can be strategically used to: Simulate donor behaviors before real campaigns, Enable low-risk testing of targeting strategies, Prepare segmentation logic for CRM integration. Going forward, enhancing synthetic data diversity will be key to unlocking its full potential for A/B testing, real-time personalization, and long-term value prediction in nonprofit fundraising.

16. Business Value of Synthetic Data

Although our current synthetic dataset was **clean and homogeneous** to fully realize the advantages of donor simulation, our work highlights the **strategic potential** of synthetic data for advancing nonprofit analytics and fundraising.

As part of the donor analytics strategy, synthetic data was leveraged to supplement gaps in the available real-world datasets. This approach provided critical support for early analysis, predictive modeling, and strategic planning in the face of data limitations.

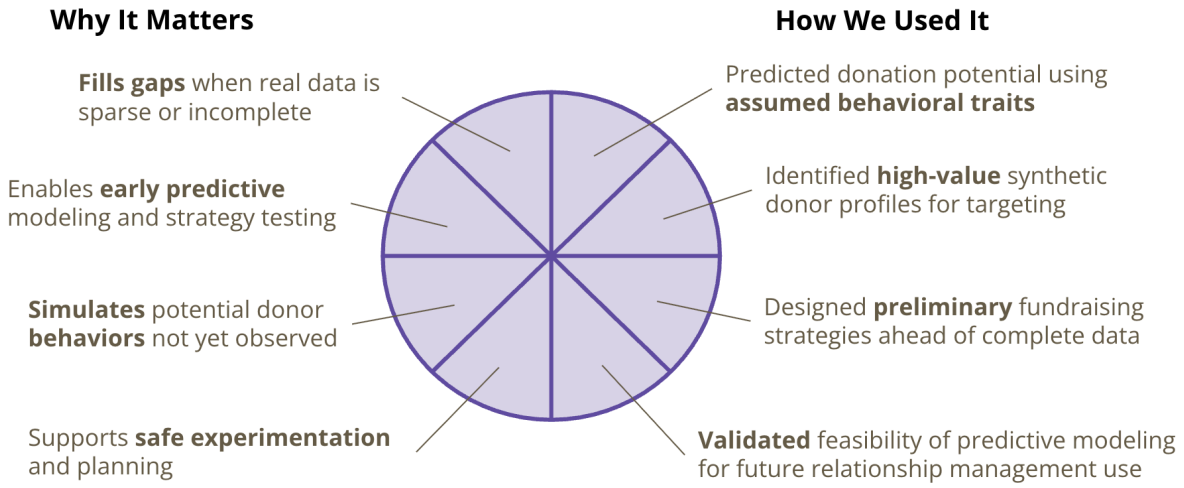
Why Synthetic Data Matters

- **Fills gaps** where real donor data is **sparse or incomplete**, ensuring a more robust dataset for exploratory modeling.
- Enables **early predictive modeling and strategic testing**, without needing to wait for multiple years of real donation behavior to accumulate.
- **Simulates behavioral patterns** not yet observed, allowing analysts to test "what-if" scenarios in donor response and segmentation.
- Facilitates **safe experimentation** with hypothetical fundraising strategies without risking real donor engagement or data misuse.

How It Was Used in This Project

- **Predicted donation potential** based on **assumed behavioral traits** of synthetic donor profiles.
- Identified **high-value synthetic donors** to explore targeting techniques prior to availability of actual high-contributing donors.
- Supported the **design of preliminary fundraising strategies**, bridging the period before complete 2024 and 2025 data was accessible.

- **Validated the feasibility** of using historical and synthetic inputs for modeling, highlighting its potential for future use in **donor relationship management systems**.



Synthetic data proved especially valuable in simulating future states, building model prototypes, and refining outreach strategy frameworks. It served not as a substitute for real data, but as a complementary tool for accelerating insights and minimizing decision delays.

17. Short-Term Strategic Plan: Maximizing Existing Resources

Building on the insights derived from behavioral segmentation, predictive modeling, and geographic analysis, this short-term strategy focuses on **region-specific actions** designed to drive immediate impact using current assets and donor intelligence.

Regional Strategic Actions

Region	Strategic Action
Phoenix	Target General Donors for **upgrade** to High Monetary or VIP via personalized campaigns.
Scottsdale & Mesa	Host **VIP engagement events** to reinforce loyalty and leverage existing high-value donors.
Flagstaff	**Activate** under-engaged high-value donors through focused outreach.
Glendale, Tempe, Chandler	**Connect** with donors using localized storytelling to build emotional bonds.
Peoria & Gilbert	**Retain** mid-tier donors through loyalty-building retention strategies.

Strategic Focus Areas

- **Upgrade**: Convert General Donors in Phoenix into higher-value tiers (High Monetary/VIP).
- **Engage**: Strengthen loyalty among VIP donors through recognition events in Scottsdale and Mesa.
- **Connect**: Build emotional resonance through community-centered storytelling in Glendale, Tempe, and Chandler.
- **Retain**: Sustain relationships with mid-tier donors in Peoria and Gilbert.
- **Activate**: Tap into newly surfaced high-potential donors in Flagstaff with targeted messaging and outreach.

These initiatives are designed for **immediate implementation** using existing donor segments and engagement infrastructure, ensuring cost-effective impact while laying the groundwork for long-term growth.

18. Long-Term Strategy: Enabling Sustainable Growth

To ensure the continuity and scalability of data-driven fundraising success, a long-term strategy has been developed around a five-stage framework: **Data Collection**, **Predictive Modeling**, **Targeted Campaigns**, **Model Optimization**, and **CRM Integration**. This strategic cycle supports adaptive fundraising through automation, learning, and integration with organizational systems.

Strategic Framework for Sustainable Growth

- **Data Collection**
  Ongoing acquisition of donation history, demographic profiles, behavioral signals, and campaign outcomes. This forms the foundation for effective modeling and segmentation.
- **Predictive Modeling**
  Leverage advanced machine learning techniques such as Random Forests or ensemble models to predict donor behavior, value, and churn risk with high accuracy.
- **Targeted Campaigns**
  Use model outputs to design personalized and scalable campaigns. Strategies will be driven by segmentation logic, predictive scores, and behavioral insights.
- **Model Retraining & Optimization**
  Ensure continuous model improvement by retraining with new data (e.g., campaign responses, giving patterns). This supports agility in donor targeting and strategy refinement.
- **CRM Integration & Fundraising Planning**
  Embed predictive scores and donor profiles into CRM systems to guide real-time decision-making and inform quarterly or annual fundraising plans.

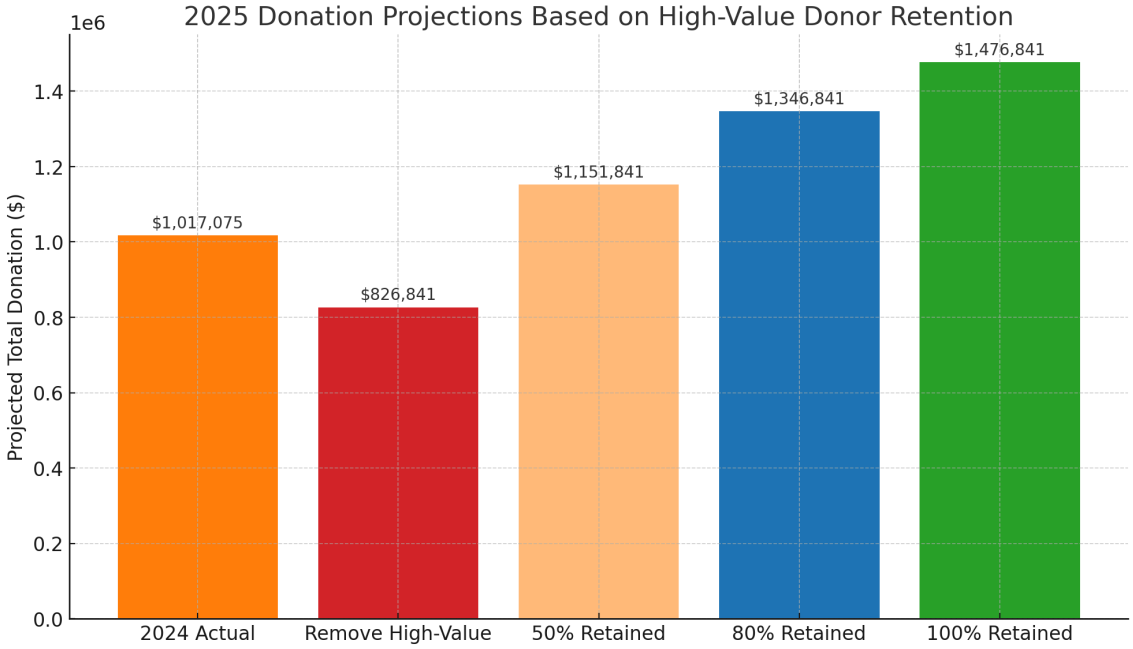
Key Components of Sustainable Growth

- **Data Enrichment**: Expand donor profiles with both internal data (RFM, engagement history) and external demographic or psychographic attributes for deeper insights.
- **Predictive Automation**: Deploy automated tools and pipelines that continuously update predictions and donor scores with incoming data.
- **Insight-Driven Campaigns**: Translate predictive outputs into actionable, audience-specific campaigns that evolve with donor behavior.
- **Continuous Learning**: Build a culture of iteration by routinely retraining and benchmarking models to maintain high performance over time.
- **CRM Integration**: Operationalize insights by embedding them within CRM workflows and aligning them with development team activities and timelines.

This approach transforms donor management from reactive to proactive, enabling long-term retention, growth, and strategic forecasting through continuous learning and system integration.

19. Impact Projection and Future Outlook

To estimate potential donations in 2025, we trained a behavioral forecasting model using donor features from 2023 and validated it by predicting actual 2024 donation outcomes. After confirming its predictive accuracy, the same model was used to project multiple donation scenarios for 2025 under varying engagement and retention strategies.



Projected Impact of Targeted Engagement

- **Boost Donations**: Prioritizing high-value donors in **Scottsdale**, **Phoenix**, and **Mesa** which together represent approximately **70% of total giving** is expected to drive the largest fundraising returns.
- **Retain Top Donors**: By applying RFM-driven loyalty strategies, the organization can preserve long-term value and prevent drop-offs among critical donor segments.
- **Improve Campaign Efficiency**: Focusing resources on **high-ROI cities** and well-defined behavioral segments ensures more efficient donor acquisition and retention per dollar spent.
- **Expand Donor Base**: Complementing high-value retention with efforts to engage **underrepresented or emerging donors** broadens the base of community support.
- **Enable Data-Driven Decisions**: All impact projections were built on **behavioral models**, reinforcing the strategic value of data-informed planning.

Forecast Scenarios for 2025

Scenario	Projected Donation	Strategic Role

2024 Actual	$1,017,075	Benchmark for performance comparison
Base Forecast	$570,783	Risk scenario assuming high-value donor loss continues
Sustained High-Value	$1,476,841	Target scenario with stable retention of current key donors
Optimistic (+10%)	$1,624,525	Stretch goal assuming improved loyalty through strategic efforts

The **bar chart** illustrates the projected donation range based on varying levels of **high-value donor retention**. Maintaining at least 80% of this segment's loyalty would lead to a projected **$1.35M+** in contributions, with further upside in optimized scenarios.

These projections clearly show that **donor retention and segmentation efforts have a direct and significant financial impact**. As such, they should remain central pillars of the organization's forward-looking fundraising strategy.

20. Conclusion

This project tackled the urgent challenge of stagnating donations faced by the Grand Canyon Council, a key regional chapter of Scouting America. By leveraging data from 2020–2024, we applied RFM analysis, clustering, and predictive modeling to uncover donor behavior patterns and inform targeted fundraising strategies.

Key Achievements:

- Donor Segmentation: Successfully classified donors into actionable RFM-based and behavior-based segments, enabling tailored outreach strategies aligned with donor value and engagement level.
- Predictive Modeling: Built a high-performing Random Forest model to forecast 2025 donations, enabling the identification of top donors and high-potential regions.
- Strategic Insights: Identified Phoenix, Scottsdale, and Mesa as critical geographies for retention and upgrade campaigns. Provided clear engagement tactics by donor type (e.g., lapsed, frequent, VIP).

Synthetic Data Integration:

- While synthetic data supported early modeling and personal development, its lack of behavioral **noise and granularity** limited its use for precise donation forecasting.

- However, it demonstrated **high potential** as a simulation tool, allowing for risk-free strategy testing and CRM-ready segmentation frameworks.

Strategic Impact:

- Our scenario projections showed that **retaining the top 20% of donors could generate $1.35M–$1.6M in 2025 donations**, emphasizing the value of segmentation and loyalty strategies.
- We recommended a **short-term plan focused on high-ROI regions** and a **long-term roadmap** involving predictive automation, model retraining, and CRM integration.

Limitations & Future Work:

Real donation data was incomplete in some years; synthetic data, while helpful, still lacks the realism needed for high-precision targeting. Future efforts should focus on: Enriching synthetic datasets with more behavioral variation, Integrating real-time feedback loops (e.g., A/B testing) into model refinement, Scaling this framework nationally across other councils.

Ultimately, by combining analytical rigor with practical strategy, this project equips the Grand Canyon Council to deepen donor relationships, expand its supporter base, and better serve the youth of Arizona for years to come.