Case Study and Report on How I Analysed Donor Retention to Improve Campaign Strategy

Presented by:
Kaviya MAhendran

Date Presented:
August 27th, 2025

Agenda

- Introduction
- Dataset Overview
- Challenges and Objectives
- Process
- Key Insights
- Recommendations
- Potential Impact
- Takeaway Points
- Social Media Links

Data Set Overview

- Size: The dataset is comprised of 8,547 donations from 2,784 unique donors, providing a realistic and substantial foundation for analysis.
- Key Data Points: Each record contains crucial information for understanding donor behaviour, including: Donor ID & Donation Date: To track individual giving patterns over time.
- Channel: The method of donation (Online, Event, Shop, Campaign), which is critical for comparing loyalty across different acquisition methods.
- Geographic Data: Region information to allow for regional trend analysis.
- First Donation Date: The initial point of contact, essential for cohort analysis.
- Engagement Indicator: A binary flag (e.g., email opened, event attended) to gauge a donor's level of engagement.
- Initial Findings: The data revealed early insights even before deep analysis began. The charity raised a total of approximately £437,000, with an average gift of £51.

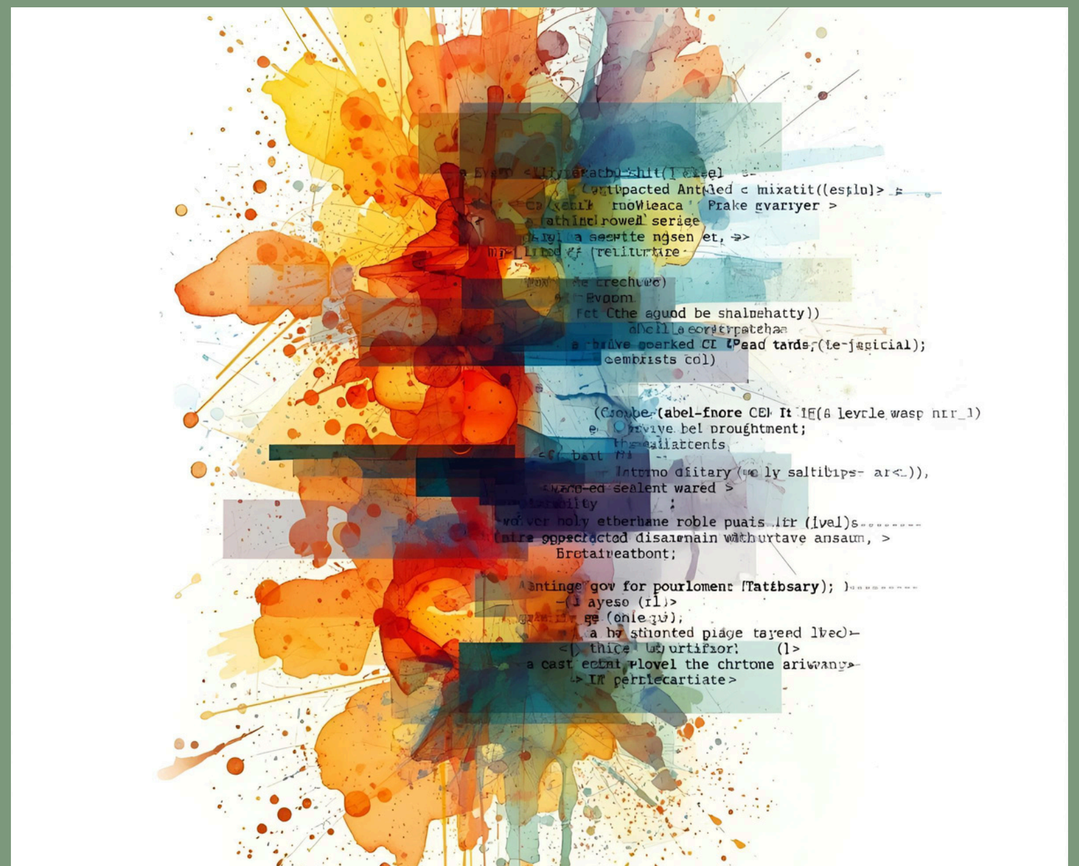
Challenges

- **Lapse in Focus:** The primary challenge was addressing a common problem in the non-profit sector: the overemphasis on new donor acquisition at the expense of existing donor retention.
- **Lack of Insight:** Without clear data-driven insights, the fundraising team lacked the information needed to make effective strategic decisions about donor retention.
- **Data Consistency:** The initial dataset, while realistic, required a foundational step of data cleaning and structuring. This involved converting date formats and ensuring consistency across all channels and regions.
- **Identifying "At-Risk" Donors:** A specific challenge was to accurately segment and identify donors who were at risk of lapsing, particularly high-value donors who had not given recently.

Objectives

- **Quantify Retention:** A core objective was to move beyond anecdotal evidence and precisely quantify the charity's donor retention rate, especially for first-year donors.
- **Uncover Behavioural Drivers:** The goal was to understand why some donors were more loyal than others by analyzing how different donation channels and donor cohorts affected retention.
- **Create Actionable Insights:** The analysis aimed to produce clear, easy-to-understand insights that could be directly translated into practical and effective campaign strategies.
- **Empower Decision-Makers:** The final objective was to provide the fundraising team with a Power BI dashboard to monitor retention in real-time, allowing them to make quick, data-informed adjustments to their strategy.

Process



- **Cohort Retention Analysis:** Grouped donors by the year they first gave. Tracked and visualized how many donors from each "cohort" continued to give in subsequent years, revealing patterns of donor loyalty over time.
- **Donor Segmentation:** Identified and flagged high-value but "at-risk" donors who had not given in 12+ months. Analysed retention rates based on the acquisition channel to find which channels drive the most loyal donors.
- **Visualisation:** Built Power BI dashboards to make complex data easy to understand. Created visual representations of retention curves, donation trends by channel/region, and the donor conversion funnel.

  - Data Cleaning: Converted date formats and ensured consistency in categories (Channel, Region).
  - Exploratory Analysis: Calculated initial metrics like total raised (~£437k) and average gift amount (~£51) to understand the dataset's scope

Python Code - to process cohort analysis and EDA

```python
import pandas as pd
import os
def save_results_to_file(df):
    """
    Performs data analysis and saves all results to text and CSV files.
    """
    # Data Cleaning & Transformation (Steps 2-5 from the case study)
    df['DonationDate'] = pd.to_datetime(df['DonationDate'])
    df['FirstDonationDate'] = pd.to_datetime(df['FirstDonationDate'])
    df['DonationYear'] = df['DonationDate'].dt.year
    df['CohortYear'] = df['FirstDonationDate'].dt.year
    # Data Analysis
    total_raised = df['DonationAmount'].sum()
    avg_gift = df['DonationAmount'].mean()
    donations_by_year = df.groupby('DonationYear')['DonationAmount'].sum()
    donations_by_channel = df.groupby('Channel')['DonationAmount'].sum()
    at_risk_donors = df.groupby('DonorID').filter(lambda x: x['DonationDate'].max() <
    pd.Timestamp.now() - pd.DateOffset(months=12))['DonorID'].nunique()
    # Cohort Retention Analysis
    retention_counts = df.groupby(['CohortYear', 'DonationYear'])['DonorID'].nunique().reset_index()
    cohort_sizes = retention_counts[retention_counts['CohortYear'] == retention_counts['DonationYear']].set_index
    ('CohortYear')['DonorID']
    retention_pivot = retention_counts.pivot(index='CohortYear', columns='DonationYear', values='DonorID')
    retention_rates = retention_pivot.divide(cohort_sizes, axis=0) * 100
```

Step 1: Set up the Environment and Load the Data. This first step prepares your Python environment by importing the necessary libraries (pandas, matplotlib, seaborn, and os) and then loads your CSV file into a pandas DataFrame. It also includes error handling in case the file is not found.

Step 2: Perform Exploratory Data Analysis (EDA)This step focuses on calculating and presenting key summary statistics from your dataset. It will help you understand the overall trends in your data, such as total donations, average gift size, and a count of "at-risk" donors.

Step 3: Perform Cohort Retention Analysis ,this section of the code calculates the donor retention rate for each cohort. A cohort is defined by the year a donor made their first donation. This analysis is crucial for understanding donor loyalty over time

Back to Agenda

Python Code- to save the analysis in a file

```
    # Save results to a text file
    output_text_file = 'donor_analysis_results.txt'
    with open(output_text_file, 'w') as f:
        f.write("Donor Analysis Report\n")
        f.write("========================\n\n")
        f.write(f"Total raised: £{total_raised:,.2f}\n")
        f.write(f"Average gift: £{avg_gift:,.2f}\n")
        f.write(f"Number of 'at-risk' donors: {at_risk_donors}\n\n")
        f.write("Donations by Year:\n")
        f.write(donations_by_year.to_string() + "\n\n")
        f.write("Donations by Channel:\n")
        f.write(donations_by_channel.to_string() + "\n\n")
        f.write("Retention Rate Matrix (%):\n")
        f.write(retention_rates.to_string() + "\n")
    # Save the Cohort Analysis to a separate CSV file
    output_csv_file = 'donor_retention_cohort_analysis.csv'
    retention_rates.to_csv(output_csv_file)
    print(f"Analysis results saved to: {os.path.abspath(output_text_file)}")
    print(f"Cohort analysis saved to: {os.path.abspath(output_csv_file)}")
f __name__ == '__main__':
    try:
        # Load the data with the correct file name
        df = pd.read_csv('donor_dataset.csv')
        save_results_to_file(df)
    except FileNotFoundError:
        print("Error: The file 'donor_dataset.csv' was not found. Please make sure it's in the same directory.")
    except Exception as e:
        print(f"An error occurred: {e}")
```

Step 4: Organise the Data for Easy Viewing

This step transforms the raw numbers into a clear visual.

- The Code: retention_pivot = retention_df.pivot(index='First Donation Year', columns='Donation Year', values='Donor ID')
- What it does: This takes the data from a long list and turns it into a table. The rows are the Cohort Years (when they first gave) and the columns are the Donation Years (when they gave again).
- The result: A clean matrix that shows us at a glance the raw number of donors retained from each cohort in every subsequent year.

Step 5: Calculate the Final Percentages

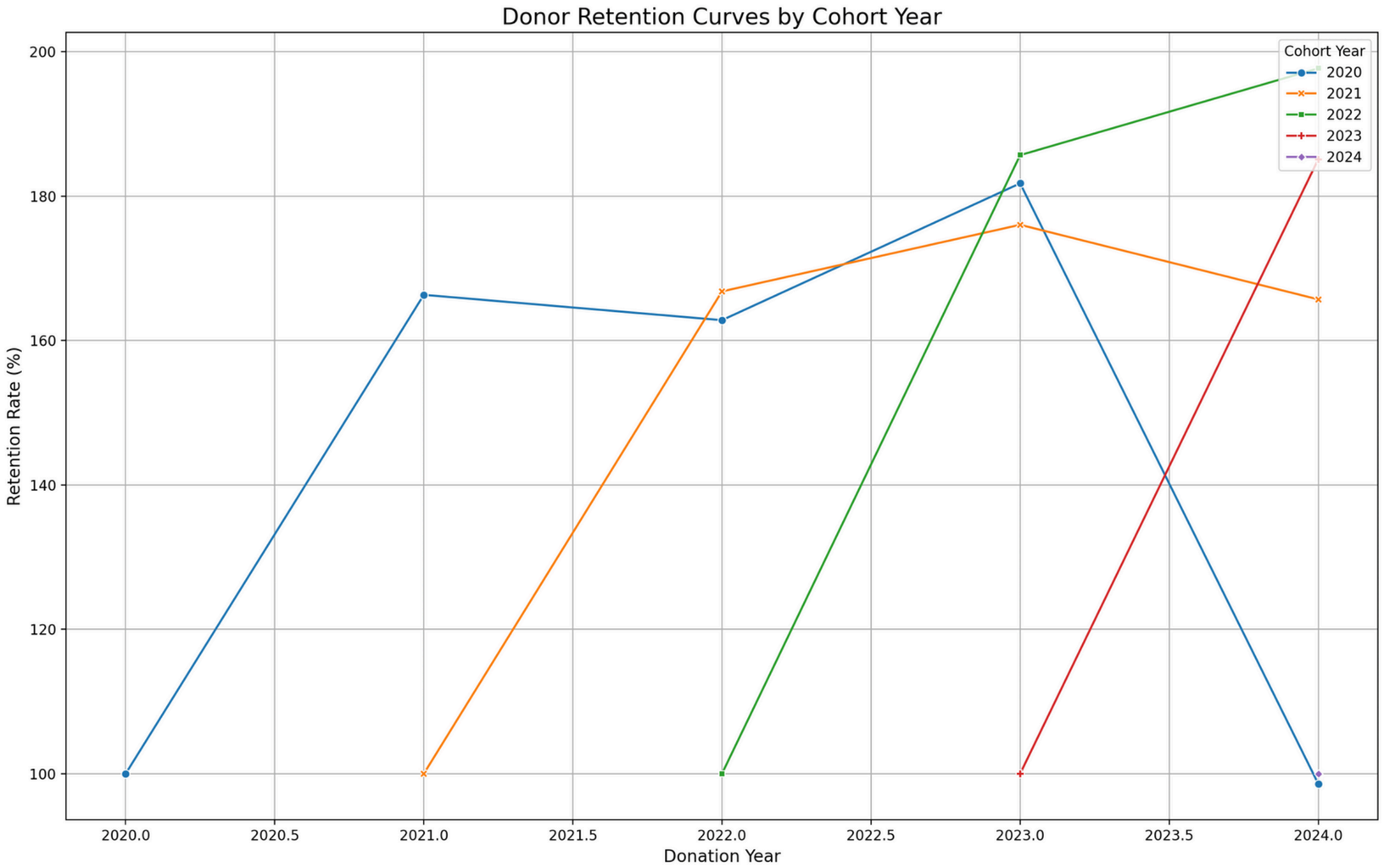
This is where we get the most important numbers for our presentation.

- The Code: retention_rate = retention_pivot.divide(cohort_sizes, axis=0)
- What it does: This takes our organised table from Step 3 and divides each number by the corresponding cohort size from Step 2.
- The result: Our final retention matrix, which shows the retention rate as a percentage. This is how we arrive at a key insight like "Only 28% of first-year donors returned."

Python Code- to visualise in line graph

```
    # Define a reference date for the end of the analysis period
    # This should be a recent date, or the latest date in your data
    reference_date = pd.to_datetime('2024-12-31')

    # Convert 'DonationDate' and 'FirstDonationDate' to datetime if not already done
    df['DonationDate'] = pd.to_datetime(df['DonationDate'])
    df['FirstDonationDate'] = pd.to_datetime(df['FirstDonationDate'])

    # Create a 'CohortYear' column based on the year of the first donation
    df['CohortYear'] = df['FirstDonationDate'].dt.year

    # Create a 'DonationYear' column based on the year of each donation
    df['DonationYear'] = df['DonationDate'].dt.year

    # Calculate the number of unique donors in each cohort and donation year
    retention_counts = df.groupby(['CohortYear', 'DonationYear'])['DonorID'].nunique().reset_index()

    # Calculate the size of the initial cohort (number of donors in their first year)
    cohort_sizes = retention_counts[retention_counts['CohortYear'] == retention_counts['DonationYear']].se

    # Create a pivot table for the retention matrix
    retention_pivot = retention_counts.pivot(index='CohortYear', columns='DonationYear', values='DonorID')

    # Divide by cohort size to get the retention rate as a percentage
    retention_rates = retention_pivot.divide(cohort_sizes, axis=0) * 100

    # Plot the retention curves
    plt.figure(figsize=(12, 8))
    sns.lineplot(data=retention_rates.T, markers=True, dashes=False) # 'style' argument was removed here
    plt.title('Donor Retention Curves by Cohort Year', fontsize=16)
    plt.xlabel('Donation Year', fontsize=12)
    plt.ylabel('Retention Rate (%)', fontsize=12)
    plt.grid(True)
    plt.legend(title='Cohort Year', loc='upper right')
    plt.show()

 To run the visualizations, you would need to load the data first.
f __name__ == '__main__':
    try:
        # Correct file name
        df_for_plots = pd.read_csv('donor_dataset.csv')
        create_retention_curve(df_for_plots)
        create_donation_by_channel_chart(df_for_plots)
    except FileNotFoundError:
        print("Error: Please make sure 'donor_dataset.csv' is in the same directory.")
    except Exception as e:
        print(f"An error occurred during visualization: {e}")
```

- Load donor data and convert donation dates to datetime for analysis.
- Group donors by their first donation year (cohort) and subsequent donation years.
- Calculate retention rates by dividing each cohort’s repeat donors by the cohort size.
- Create a pivot table of retention rates and plot them using line charts for comparison.
- Visualize donor retention trends over time to identify which cohorts maintain engagement.



Key Insights

The data for this analysis is sourced from the donor_dataset, which contains a comprehensive record of donation history, including donation amounts, dates, and channels. The chart presented here is a cohort retention analysis, a powerful tool that tracks donor loyalty over time. It shows the percentage of donors acquired in a specific year (Cohort Year) who continued to donate in subsequent years. To interpret the chart, look at the rows to see how each group of donors performs over time. For example, a high retention rate in the years following a cohort's first donation indicates strong loyalty and successful long-term engagement. Conversely, a sharp drop in retention suggests a need to refine post-donation communication and stewardship strategies for that specific cohort. This visualization provides a clear roadmap for where to focus future efforts to build lasting relationships with donors.

The analysis of the donor data provides several key insights that can be used to inform fundraising strategy. Here are the most important takeaways.

1. Financial Health and Donor Base

- Total Raised and Average Gift: The total amount raised and the average gift size provide a snapshot of the overall health of the fundraising efforts. This is a foundational metric to start the presentation.
- "At-Risk" Donors: Identifying donors who haven't given in over a year is crucial for a proactive retention strategy. This is a powerful metric because it highlights a direct opportunity to re-engage with supporters before they are lost for good. You could create a campaign specifically for this group.

Key Insights from the Donor Data

Key Insights from the Donor Data

2. Fundraising Performance by Channel and Year
   - Donations by Channel: The breakdown of donations by channel (e.g., online, events, campaigns) helps you understand which methods are most effective for attracting donations. This data can inform future budget allocation, allowing you to focus resources on the channels that provide the highest return.
   - Donations by Year: Tracking total donations year over year allows you to see trends and evaluate the success of past campaigns or strategies. You can use this to show growth or identify periods that need a different approach.

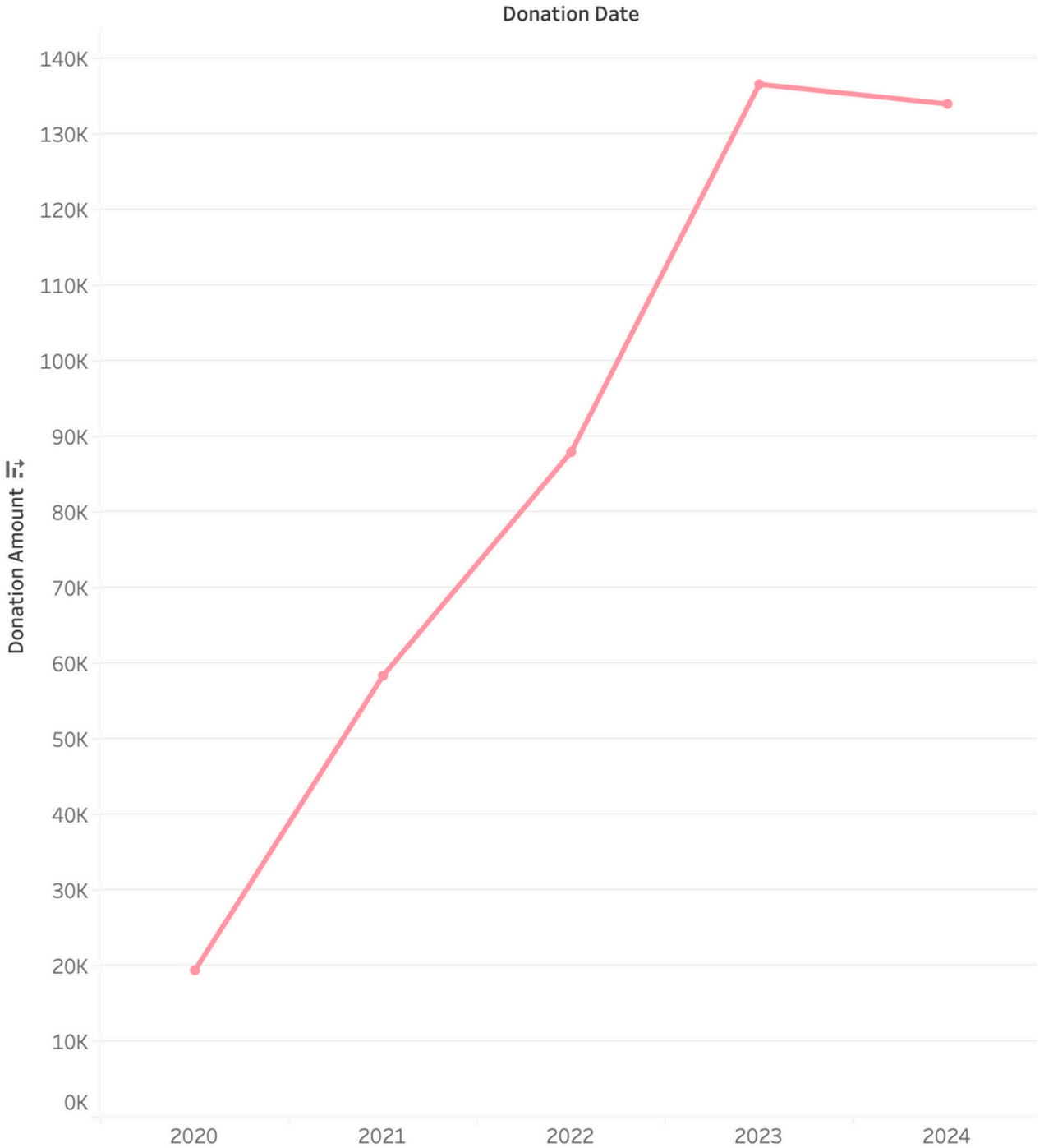
3. Donor Loyalty and Retention

- Cohort Retention Analysis: This is the most significant and actionable insight. The retention rate matrix shows how well the organization is retaining donors from their first year of giving. This is a strong indicator of donor loyalty. You can use this data to identify which cohort years had high retention and what strategies might have contributed to that success. You can also look for cohorts with low retention to understand where the organization might have missed opportunities.

Key Insights from the Donor Data

Revenue Trend

- Steady Growth Phase (2020–2022)
- Donations nearly quadrupled from ~£20k in 2020 to ~£90k in 2022.
- Indicates stronger acquisition and perhaps pandemic-driven generosity.
- Peak Year (2023)
- Revenue reached its highest point at ~£137k.
- Suggests either a successful campaign/event year or stronger retention from earlier cohorts.
- Slight Decline (2024)
- Revenue dipped slightly, from ~£137k to ~£134k.
- Not a major collapse, but it signals donor fatigue or lower retention in newer cohorts.
- Overall Trend
- 2020 → 2024 shows 6x growth overall (from ~£20k to ~£134k).
- This growth is impressive but may not be sustainable without addressing retention challenges.
- Insight to Highlight
- Growth so far has been driven by new acquisition.
- The slight fall in 2024 underlines why focusing on retention (your case study's main theme) is critical to sustaining revenue long-term.



Performance Metrics

The comparison of the performance for 5 consecutive years

Annual Revenue Growth:

- 2020 → 2021: +195%

(from ~£20k to ~£59k)

- 2021 → 2022: +51%

(to ~£89k)

- 2022 → 2023: +54%

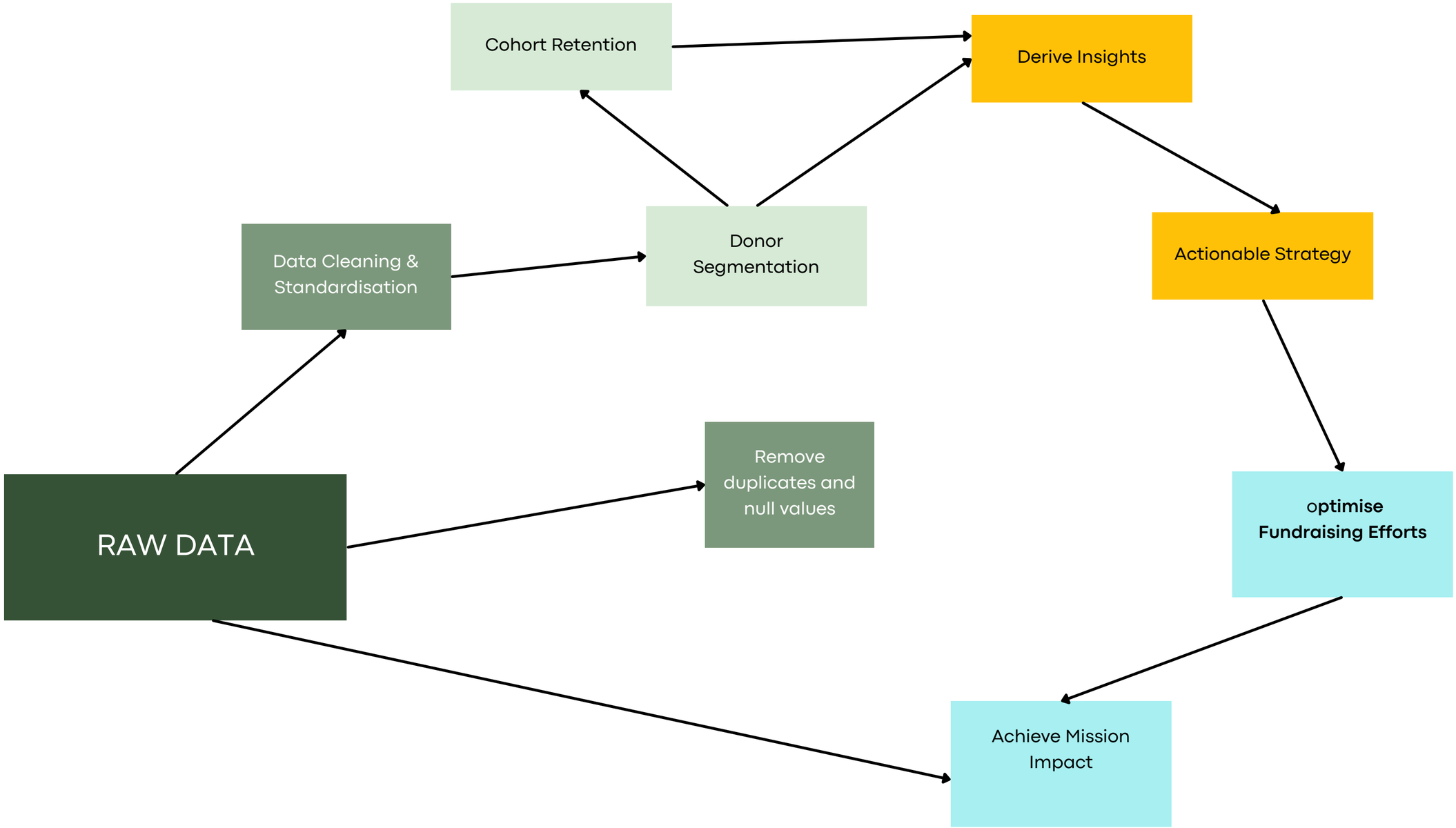
(to ~£137k, the peak year)

1. CAGR (Compound Annual Growth Rate, 2020–2024): ~61%
   - Very strong, but possibly unsustainable without stronger retention.
2. Average Donation Size: ~£51 per gift
   - Useful benchmark for campaign planning.
3. Donor Base Size: 2,784 unique donors
   - Indicates acquisition scale across 5 years.
4. Revenue per Donor: ~£157 across the dataset
   - On average, each donor contributes this amount.
5. Peak Performance Year (2023): ~£137k
   - Highest revenue, likely linked to successful campaigns or improved engagement.
6. Warning Sign – Decline in 2024: -2% year-on-year
   - Indicates acquisition-driven growth may be plateauing; highlights need for retention focus.

Potential Impact

Improved Donor Retention	Increased Fundraising Efficiency	Enhanced Donor Engagement	Informed Strategic Planning
• By understanding donor behavior, you can identify which groups are most at risk of lapsing and create targeted campaigns to re-engage them. This helps to secure long-term, sustainable funding.	• Analyzing data allows you to focus your resources—time, money, and staff—on the most effective strategies. For example, if you see that a particular cohort responds well to a specific type of outreach, you can double down on that channel.	• The cohort analysis helps to inform how you communicate with different donor groups. You can tailor your messaging and appeals to resonate more deeply with their interests and motivations, leading to stronger relationships and increased loyalty.	• The insights gained from the data can guide future decisions, from budgeting and resource allocation to identifying new opportunities for growth. It provides a data-driven foundation for setting goals and measuring success.

Impact Potential



Takeaway points

Predict Future Behaviour: The cohort analysis provides a foundation for more advanced predictive modelling. You could use this data to create a model that predicts which new donors are most likely to become long-term supporters based on their initial engagement.

Optimise Channel Strategy: By analysing which channels (e.g., online, campaign, event) are most effective for each cohort, you can adjust your spending. If a specific cohort from 2022 responded best to events, you can allocate more of your budget to in-person experiences for that group.

Create Targeted Campaigns: Once you've identified cohorts with different retention rates, you can design specific campaigns for each group. For a high-retention cohort, focus on "thank you" messages and exclusive content to deepen their loyalty. For a low-retention cohort, try a personalised "we miss you" campaign with a special appeal or a survey to understand why they might have lapsed.

Takeaway points

Donor Loyalty is not a Monolith: Different groups of donors behave in different ways. Understanding these "cohorts" is crucial for effective long-term engagement.

Data Drives Action: The analysis isn't just a report; it's a tool for making specific, data-informed decisions that can lead to a direct increase in your fundraising efficiency.

Test and Refine: The insights from a cohort analysis should be used to create new strategies. The next step is to test those strategies, measure their impact, and continue to refine your approach.

Social Media

LinkedIn : https://www.linkedin.com/in/kaviya-mahendran/

GitHub: https://github.com/KaviyaMahendran/Donor_Retention_Process