I. Problem description

1. Background

The United Nations Children's Fund (UNICEF) is one of the world's leading humanitarian organizations, operating in over 190 countries and territories to protect the rights and well-being of every child. Through programs in health, education, emergency relief, and child protection, UNICEF works to ensure that the most vulnerable children have access to life-saving services and long-term support.

Although UNICEF is globally funded by UN member states and public donations, its funding structure remains vulnerable to external changes. For example, recent announcements forecast a projected 20% reduction in UNICEF global funding by 2026, largely due to decreased contributions from major donors such as the US (Reuters, 2025). As a result, UNICEF faces increasing pressure to run its fundraising campaigns more efficiently. Identifying stable revenue sources has become crucial to ensuring the sustainability of its mission.

2. Problem description

2.1 Key challenges

- **Unstable income stream:** Due to the voluntary nature of donations, fluctuating funding is a sector-wise challenge triggered by external factors such as policy shift or economic downturns
- **Reliance on one-off donors:** One-off donations are inconsistent and often driven by specific events or emotional appeals. In fact, half of single donors never return to give again, making long-term planning for ongoing programs difficult (Bagot et al., 2016).
- **High fundraising cost and transparency requirements:** Attracting new donors requires significant investment in marketing and administrative processes. When donor retention is low, these acquisition costs reduce overall fundraising efficiency and dilute resources for direct impact. At the same time, charities are expected to maintain financial transparency, adding operational pressure (ACNC, 2021)
- **Lack of Predictable Revenue:** While emergency campaigns can provide short-term funding relief, they are unpredictable, challenging income forecast and limit the ability to commit to long-term projects
- **Limited data-driven targeting:** There is significant room for improvement in leveraging data to better identify and engage supporters who are most likely to commit to long-term giving.

2.2 Strategic approaches and actions

- **Strengthen the regular giver base:** Regular givers offer a predictable and recurring revenue stream, offering greater financial stability. Over time, they tend to deliver higher cumulative value than one-time donors, making them a more profitable and sustainable supporter segment.
- **Refine acquisition and retention strategy:** Shift focus from mass outreach to building long-term donor relationships. This can be achieved by improving communication strategies that encourage engagement beyond the first donation.

- **Strengthen Donor Segmentation:** Develop clearer supporter personas or audience groups to allow more personalized and targeted communication.

3. Importance of success

Successfully enhancing the regular giver base holds significant value for UNICEF's long-term sustainability.

- Enabling more effective program planning and resource allocation
- Reducing reliance on constant donor acquisition, lowering fundraising costs over time
- Delivering consistent, life-changing support to children in need both in response to emergencies and through long-term development initiatives.

II. Data Pre-processing

1. Initial preparation for data joining

- Variables such as "Age_Bucket" and "Content_Thematic_Area", containing over 90% missing values, were excluded from the dataset due to their limited usefulness.
- "Mailing_Country" is also dropped as all entries are duplicated variations of "Australia," offering no meaningful insights. Similarly, 'Mailing_City' is excluded due to its limited relevance and overlap with "Mailing_postcode", which is used to link with external datasets. The 'ParentID' column is removed, as it cannot serve as a linking key between datasets and lacks little analytical value.

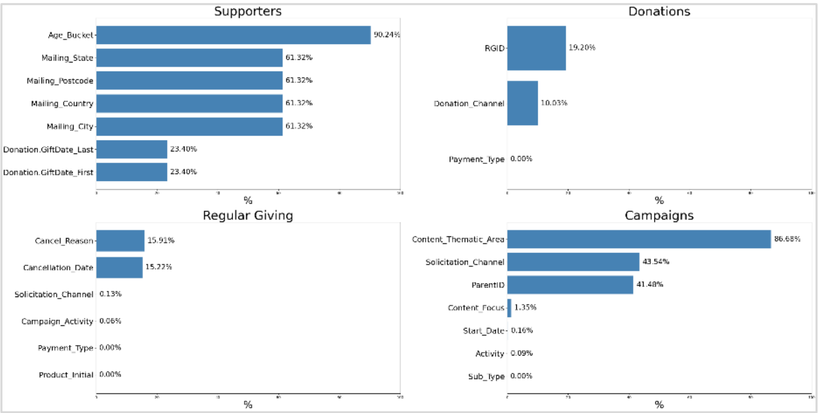


Figure 1: Proportion of missing values by datasets

- All 5 datasets contribute relevant information, so variables with moderate levels of missing data are retained at this stage to allow for data joining. This approach helps avoid inconsistencies or misalignment issues during later data cleaning steps.

2. Feature Engineering and Data Joining

2.1 Core variables to retain

- Final data set is created by joining new variables with selected columns retained from Supporters dataset. Key retained columns include:
  - Supporters data: *SupporterID, Donation.GiftDate_First, Donation.GiftDate_Last, Donor_Type, Mailing_Postcode, Mailing_State*, and all variables related to communication preferences such as *Have_HomePhone, DoNotContact, DoNotPhone, Have_Email, DoNotSMS,* etc.
  - Regular giving data: *Product_initial* and *Product_current*

2.2 Data joining from existing data

- **Giver types: Regular vs Non-regular**

- This variable “Regular_giver” is set to True if the supporter’s ID appears in the Regular Giving table.
- While some donors who are not registered as regular givers may still make multiple donations, their giving patterns are typically inconsistent and unpredictable. Unlike those who are officially registered with a fixed donation frequency (e.g., weekly or monthly), these donors are less likely to provide a stable source of recurring revenue.
- Therefore, restricting this variable to only those listed in the Regular Giving database supports the goal of forecasting conversion probability, while avoiding bias introduced by irregular giving behavior.

- **Whether this regular giver converted from a one-off giver**

- “Converted_to_regular” is set to True if the supporter has no RGID in “Donations” but has one in “Regular Giving”, suggesting a conversion from a single to a regular giver.

- “Avg_Donation_Amount”, “Total_Donation_Amount”, “Donation_Count”: These variables are derived from the Donations table by aggregating donation amounts and counts per supporter.

2.3 Data joining from external data

- **MOSAIC data**

By matching “Mailing_Postcode”, each supporter was assigned a MOSAIC group reflecting their household lifestyle, living area, income level, etc. This segmentation provides marketing-relevant insights into preferences, values, and socioeconomic status.

- **ABS data: SEIFA (Appendix B)**

To enrich supporter profiles, two indexes from the ABS SEIFA dataset were merged using the postcode data

- **IRSAD – Index of Relative Socio-Economic Advantage and Disadvantage:** This index captures both relative advantage and disadvantage by summarizing the economic and social conditions of households within a postcode area. A low decile score (1–3) indicates relative disadvantage, while higher scores (8–10) reflect greater socio-economic advantage (ABS, 2023).
- **IEO – Index of Education and Occupation:** This index focuses solely on education attainment and occupational status (ABS, 2023). A low score reflects areas with fewer qualifications, lower-skilled jobs, and higher unemployment while a high score indicates communities with higher education levels and skilled occupations. Unlike IRSAD, the IEO omits income data, providing a more focused lens on human capital indicators.

3. Data cleaning

- Key variables including “Mailing_Postcode” and “Donation.GiftDate_Last” and “Donation.GiftDate_First” are essential for linking to external datasets and identifying target geographic areas. Rows with missing values in key variables are not suitable for analysis since they limit the ability to generate meaningful insights. Therefore, despite the relatively high proportion of missing data, all rows with missing values in these fields are removed from the dataset.

- Sample sizes ranging from 1000 to 10000 are generally considered sufficient for predictive modelling and hypothesis testing (Singh & Masuku, 2014). Although the data cleaning process results in a reduction of the original sample size to **225332 supporters**, the resulting dataset remains **substantially large** for robust statistical analysis and modelling.
- Moreover, the excluded data contains missing values in critical fields, making them unsuitable for comparative or predictive analysis. Retaining them could compromise the validity of results or introduce bias due to incomplete or inconsistent features (Kuhn & Johnson, 2013).
- Beyond missing values from the original datasets, some missing data are introduced due to misalignment between datasets. For instance, certain supporters listed in the Supporter dataset have no recorded donations in the Donations dataset. Similarly, some regular givers in the Regular Giving dataset do not have a matching Supporter ID in the Supporter dataset. However, these unmatched data represent a small proportion of the total sample, ranging from 0.45% to 4.75% (Appendix C), and are removed without significantly impacting the overall dataset size.

The **final sample size** consists of **213605 rows**

4. Data type and Data Transformation

4.1 Data type

- All categorical variables with more than 2 categories are ordinal and will be retained in their original format. This approach supports interpretation during EDA and visualization as category labels are more intuitive than numerical encodings. Dummy variables may be introduced later during modelling stages if required.
- Binary or nominal categorical variables have been converted into numeric format for analysis. Values such as “Yes” or “True” are assigned a value of 1, while “No” or “False” are assigned 0.
- All numerical variables are discrete and are presented in appropriate format

4.2 Data transformation

Both the average gift amount and total gift amount exhibit strong positive skewness due to the presence of large donation outliers (fig.4). To normalize these distributions and reduce skewness, the natural logarithm transformation is applied to both variables (fig.2, fig.3)

Figure 2

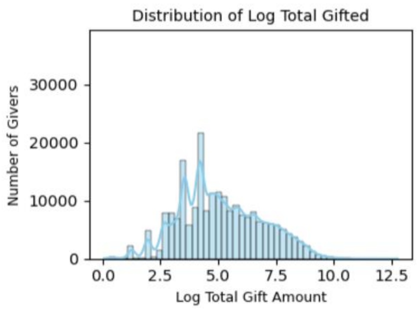


Figure 3

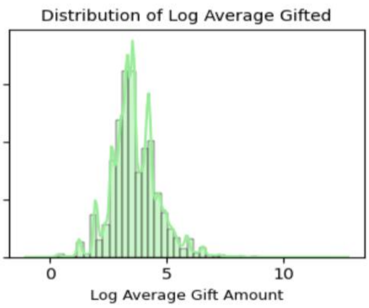


Figure 4: Donation amount statistics

	Avg_Gift_Amount	Total_Gift_Amount
Skew	323.51	38.64
Kurt	120001.98	3779.04

III. Descriptive statistics

1. Overall donors analysis

Key observations

- Almost 76,400, representing 35.8% of the sample size, are classified as regular givers. The remaining 137,237 donors (64.2%) have either donated only once or have not registered for recurring giving. Overall, the majority of the donor base consists of non-regular givers.
- Among the regular givers, only 12,834 donors (16.8%) were identified as having converted from one-time to recurring donors. The remaining 63,534 regular givers (83.2%) are assumed to have either signed up directly as recurring donors or adopted that behaviour prior to the period covered by this dataset.
- While a considerable number of donors are classified as regular givers, only a small fraction have transitioned from one-time to regular giving, suggesting that most regular donors are acquired directly rather than through conversion.

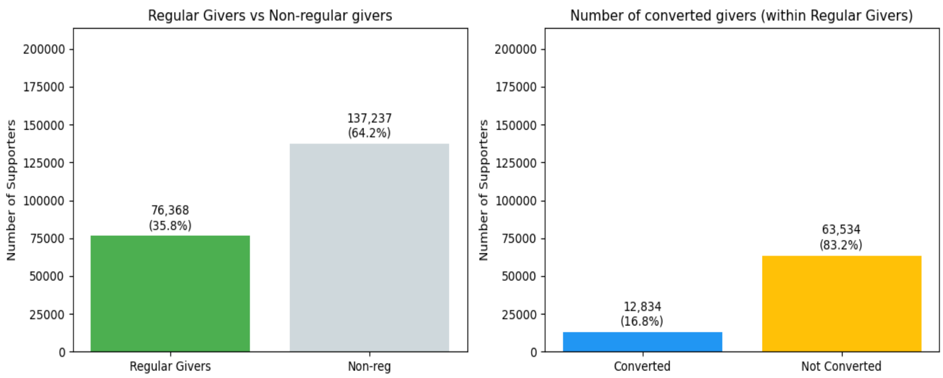


Figure 5: Giver categories breakdown

2. Key variables distributions

2.1 Donation amount features

- The distributions of both average and total donation amounts are highly right-skewed, with extreme skewness and kurtosis values (fig.5). This indicates a long right tail, where most donors contribute small amounts, while a few give extremely large sums

Statistics Summary

Figure 6: All givers

	Avg_Gift_Amount	Total_Gift_Amount
count	213605	213605
mean	76.87	817.59
std	887.92	2691.01
min	0.35	1.05
max	350000	350000
Skew	323.51	38.64
Kurt	120001.98	3779.04

Figure 7: Regular givers

	Avg_Gift_Amount	Total_Gift_Amount
count	76368	76368
mean	28.73	1706.96
std	32.16	3315.27
min	0.35	1.40
max	1808.33	256097.29
Skew	323.51	38.64
Kurt	120001.98	3779.04

Figure 8: Non-regular givers

	Avg_Gift_Amount	Total_Gift_Amount
count	137237	137237
mean	103.65	322.69
std	1106.60	2114.25
min	0.62	1.05
max	350000	350000
Skew	323.51	38.64
Kurt	120001.98	3779.04

- In particular, the majority of supporters have total donation amounts under $100. This is largely due to the high proportion of one-off givers, whose single contributions keep total values concentrated near the lower end of the scale.
- A small number of high-value donors account for the long right tail. While these outliers heavily influence the mean and standard deviation, they are retained in the analysis because they represent the group we aim to understand, high-value donors. Though they make up a small fraction of the donor base, they contribute a large share of total revenue.

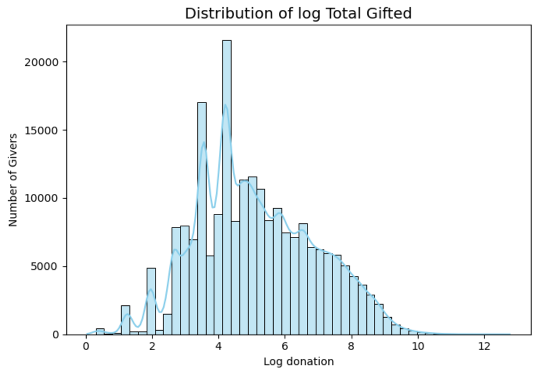


Figure 9

- To enable meaningful analysis across all donor levels, supporters are grouped into five quartiles using “qcut", based on their total donation amounts.

2.2 Donation count

- The distribution of Donation_count is positively skewed, indicating that while most supporters donated only a few times, a small number contributed very frequently. The median value of 2 suggests that at least half of all supporters donated twice or less. Additionally, 75% of donors made 16 or fewer donations.

Figure 10

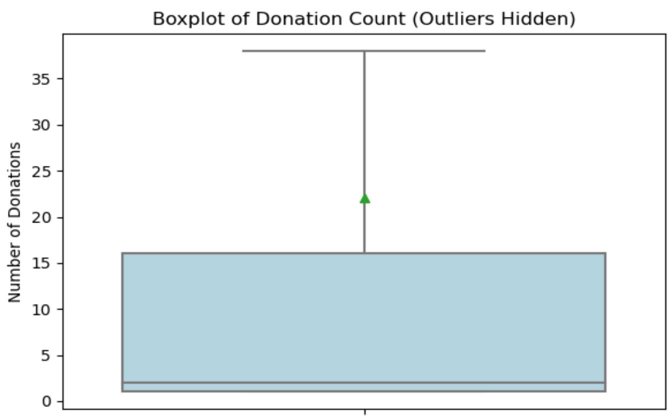


Figure 11: Donation count statistics

	Donation_Count
count	213605
mean	22.07
std	46.78
min	1
25%	1
50%	2
75%	16
max	878
Skew	3.31
Kurt	13.11

- The maximum donation count reaches 878, though such extreme values are rare and likely represent regular givers who contribute on a fixed schedule. The interquartile range, ranging from 1 to 16 donations, reflects a wide variation within the typical donor population.

3. <u>RFM features construction</u>

To support UNICEF’s goal of identifying one-off supporters who are likely to convert into regular givers, RFM analysis is conducted by transforming key variables in the dataset into behavioural indicators of donors. RFM (Recency, Frequency, Monetary) is a framework that captures how recently a person donated, how often they donate, and how much they give (Christy et al., 2021). These three factors are important predictors of future engagement and conversion potential.

- **Recency** was calculated by measuring the number of days since each supporter's most recent donation, using 1 day after the latest donation date in the dataset as the cut-off. Donors were then divided into five equal-sized groups, with the most recent donors receiving the highest recency scores.
- **Frequency** was based on the total number of donations made by each supporter. Frequent donors are more likely to adopt regular giving because they have already demonstrated a pattern of repeated support. Supporters were ranked into five frequency levels, from least frequent to most frequent.
- **Monetary** value was measured by calculating the total amount each supporter had donated. This reflects their overall contribution to UNICEF's campaigns. Supporters were again divided into five groups, with those contributing the highest amounts receiving the top scores.

By integrating these three dimensions, each supporter received an RFM score that highlights their engagement level, consistency and contribution value. These scores are used to segment the supporter base and identify individuals with high potential for regular giving campaigns, helping UNICEF optimise its outreach and maximise long-term donor returns.

IV. Potential relationships investigation

4. Correlations matrix

To identify potential relationships between key variables and regular giving or conversion likelihood, a correlation matrix was generated. This matrix uncovers patterns that can inform feature selection for modelling and strategic donor targeting.

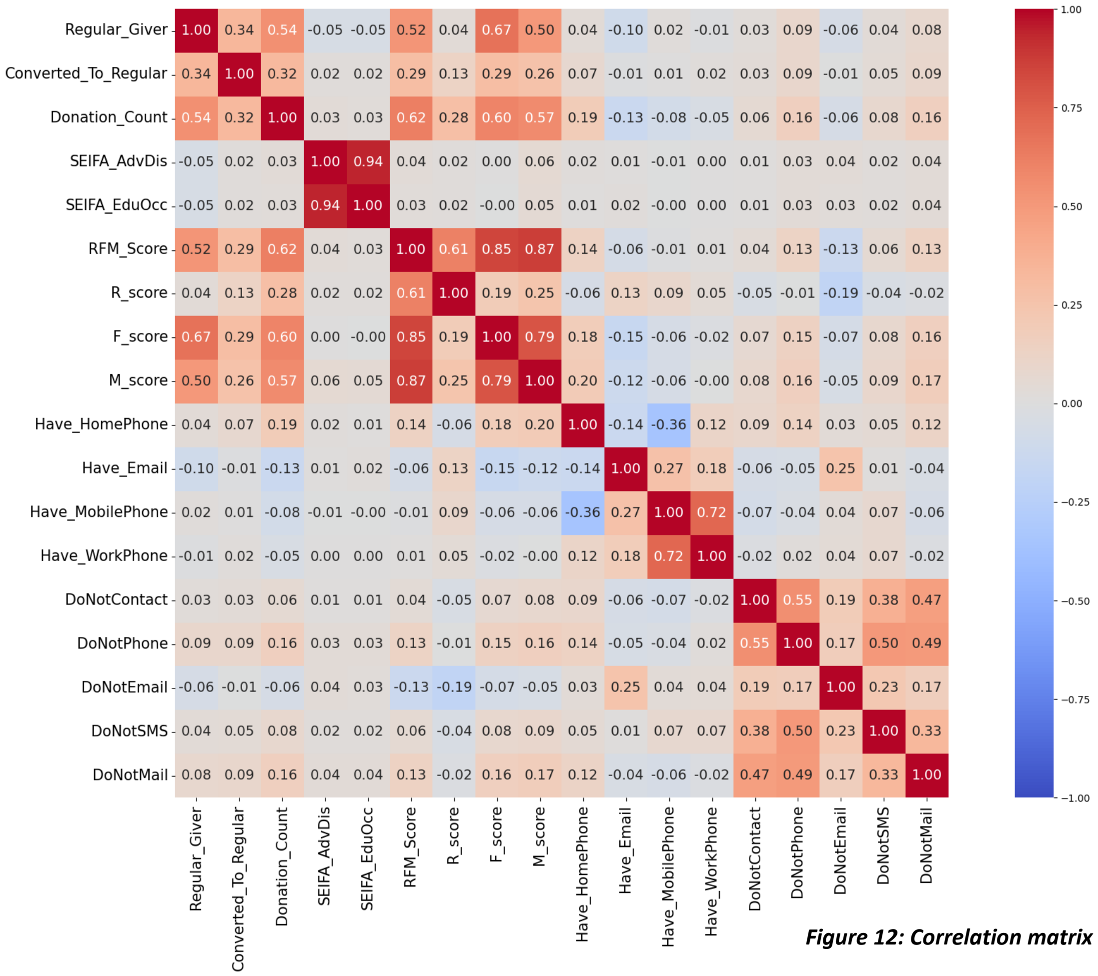


Figure 12: Correlation matrix

The top correlated variables with **Regular_giver** are: (1) Frequency score: 0.67, (2) Donation count: 0.54, (3) RFM score: 0.52, (4) Monetary score: 0.50

The top correlated variables with **Converted_To_Regular** are: (1) Donation count: 0.32, (2) RFM score: 0.29, (3) Frequency score: 0.29, (4) Monetary score: 0.26

Key observations

- **Donation_Count** exhibits a strong positive correlation with both **Regular_Giver** and **Converted_To_Regular**, which makes sense since donors registered on regular giving are committed to recurring contributions resulting in a higher number of donations
- **RFM_Score** and its components, particularly **F_score** and **M_score** are highly correlated with regular giving, reinforcing the usefulness of RFM analysis in predicting donor behavior. These results suggest that donation frequency and monetary value are key predictors of regular giving potential.
- Interestingly, variables related to **SEIFA indexes** show weak correlations with **Regular_Giver**. This is somewhat unexpected, as higher socioeconomic status is typically associated with more consistent engagement.
- Other features related to communication channels like **Have_Email** and **DoNotEmail** do show slight relationships with giving behaviours, they are comparatively minor.

5. <u>Donation amount and giver type</u>

- Total donation amounts are significantly higher for regular givers, indicating that they contribute more cumulatively. However, non-regular givers tend to give more per donation (fig.12).
- Regular givers donate in a more consistent pattern, with points closely clustered around the trendline. In contrast, non-regular givers exhibit more variability, while converted givers fall in between, reflecting a transition in donation behaviour (fig.13)

Figure 13

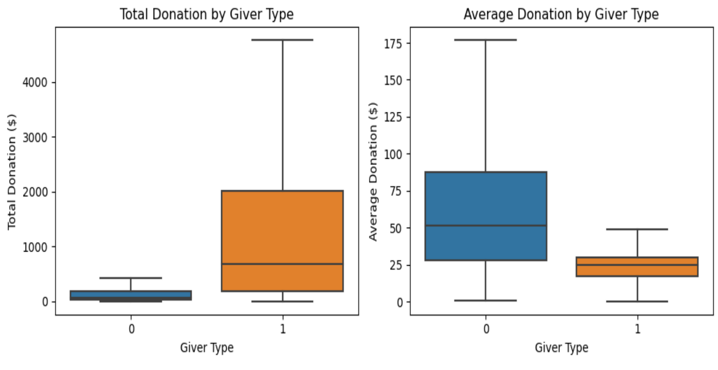
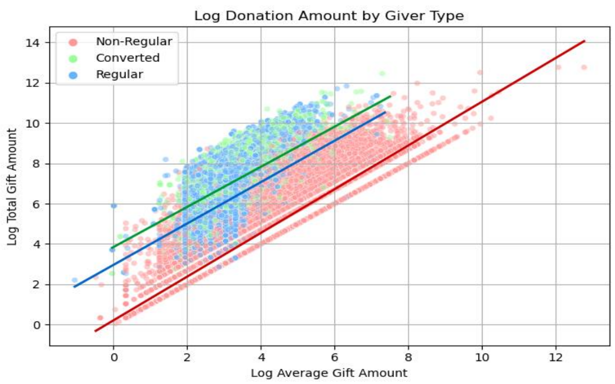


Figure 14



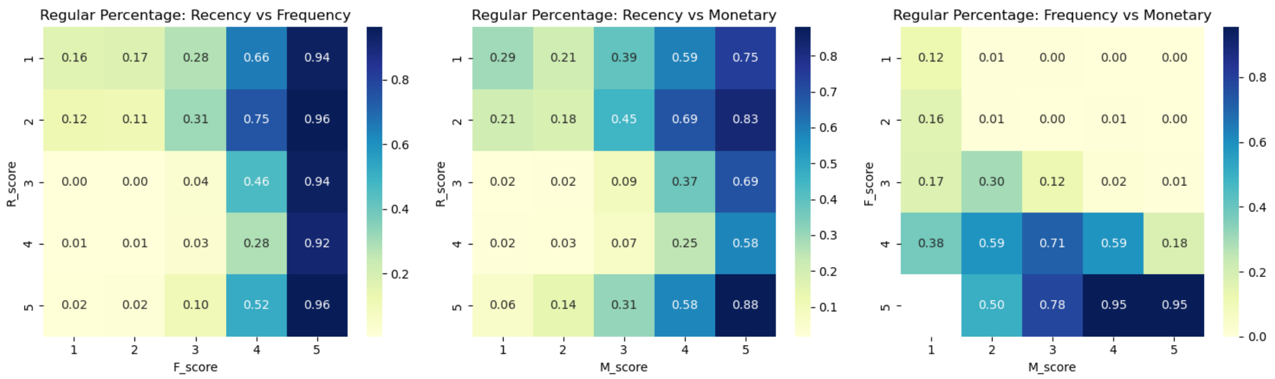
6. <u>Giver type & conversion rate and RMF score</u>

3.1 RFM factors and Proportion of regular giver

Three heat maps show how combinations of two RFM dimensions (Recency - Frequency, Recency - Monetary, Frequency - Monetary) correlate with the proportion of regular givers.

Key observations

Figure 15: Pair heatmaps between Recency, Frequency and Monetary



- **Recency - Frequency:** The proportion of regular givers increases most significantly with higher frequency scores, particularly when paired with recent giving (high R score). However, some regular givers appear in lower recency and frequency groups, likely due to **instalment patterns** such as semi-annual or annual giving, which result in longer donation intervals and lower frequency despite their regular status.
- **Recency - Monetary:** Regular giving rates appear highest not only for recent and high-value segments, but notably also for those with low recency) and high monetary scores. This reflects donors who give large amounts less frequently, such as through annual instalments
- **Frequency - Monetary:** The number of regular givers peaks when both frequency and monetary scores are high, with frequency showing a stronger overall impact.

3.2 Regular Giver Rate & Conversion Rate by RFM Score

This plot shows the percentage of regular givers and converted donors across RFM scores ranging from 3 to 15.

Key observations

- Both proportion of regular givers and conversion rate increase steadily with higher RFM scores. Conversion remains relatively low across all scores but increases consistently from RFM = 7 onwards. In contrast, the regular giver rate increases more sharply, with a significant rise starting around 10 and peaking at 15.
- Notably, a substantial number of regular givers are still found at the lower end of the RFM scale. This can be explained by the composition of the RFM score: regular givers typically donate more frequently and recently, earning high F and R scores. However, they often contribute small amounts per donation, especially those on weekly or monthly giving plans, which results in a lower M

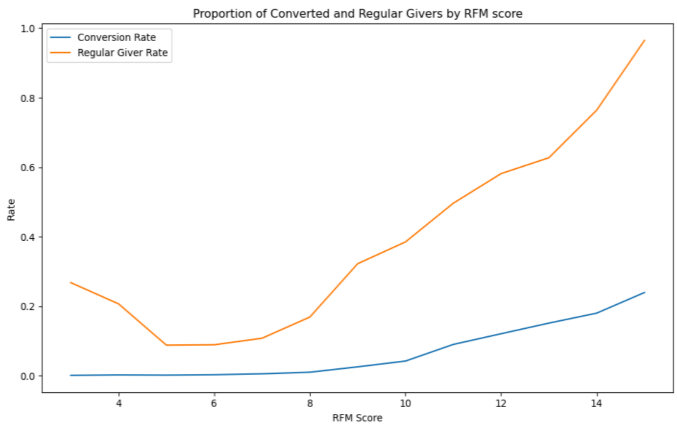


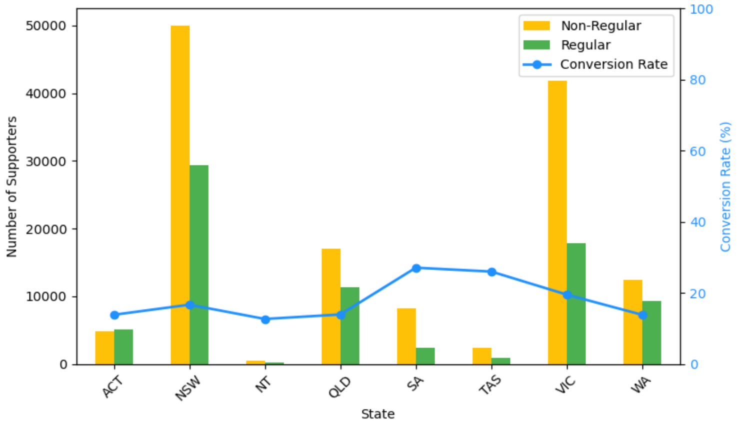
Figure 16

score. The low M score can offset the high frequency and recency scores, pulling the total RFM score into the lower range.

7. Giver type & conversion rate by states

Key observations

Figure 17: Giver type and conversion rate by state



- NSW and VIC have the largest number of supporters, with a large base of non-regular givers. In contrast, regular givers are more evenly distributed across VIC, NSW, and QLD.
- Interestingly, states like SA and TAS have fewer total supporters but exhibit the highest conversion rates, both exceeding 30%. This suggests potentially more effective donor engagement strategies in these regions. Meanwhile, despite their large supporter bases, NSW and QLD display lower conversion rates, indicating room for improvement in converting one-time donors.

8. Product distribution among regular givers

Figure 18: Current product distribution of regular givers

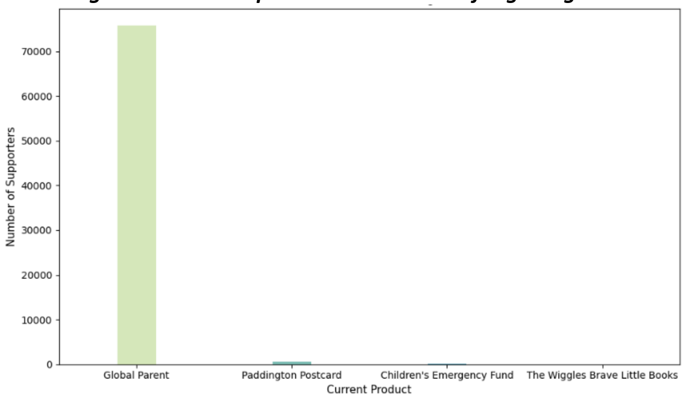
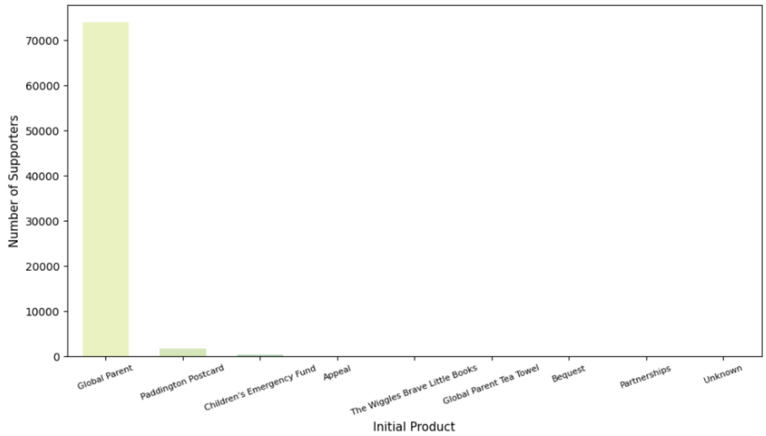


Figure 19: Initial product distribution of regular givers



Key observations

The majority of regular givers were initially acquired through the **Global Parent** campaign. Furthermore, the comparison between initial and current products shows very little movement across different product types. Most regular supporters have remained with the same product they started with, particularly Global Parent. Other products like **Paddington Postcard**, **Children's Emergency Fund**, and **Global Parent Tea Towel** account for only a small fraction of regular givers, both initially and currently.

9. Conversion rate by MOSAIC groups

Chi-Squared Test of Independence: Mosaic_Group and Regular_Giver

H_0: Mosaic_Group and Regular_Giver are independent

H_1: Mosaic_Group and Regular_Giver are related

A significance level of α =0.05 is used. The test produced the following Chi-squared statistics:

- Chi-squared statistics: 1306.55
- p-value: 2.06×10^{-271}

Since the p-value is far below 0.05, we reject the null hypothesis and conclude that Mosaic_Group and Regular_Giver are significantly related. The observed and expected frequency tables used in this test are provided in Appendix D.

Key observations

Figure 20

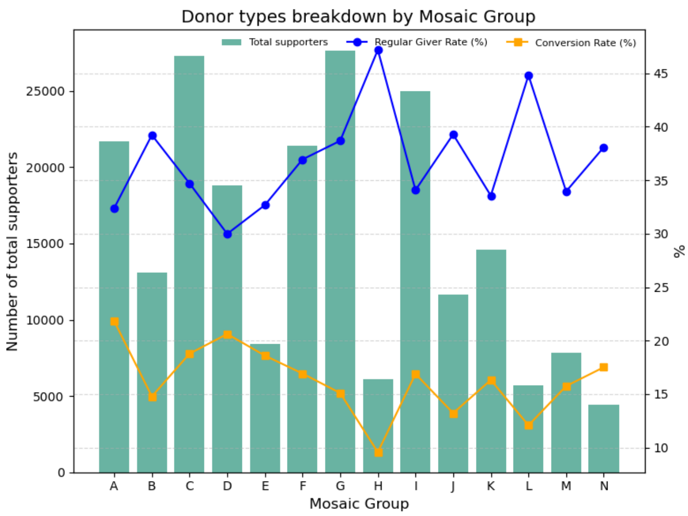
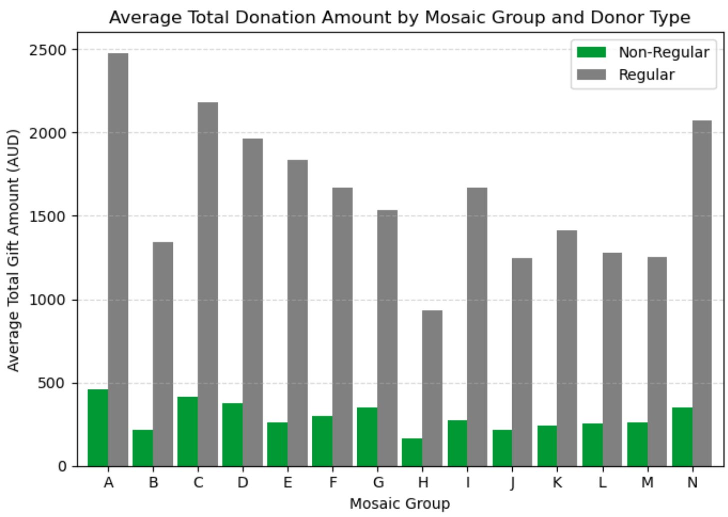


Figure 21



- Groups G (Growing Independence), C (Striving for Status), and I (Traditional Pursuits) have the highest number of supporters. These segments typically consist of financially stable, middle-aged families with average to high incomes, living in both metropolitan and suburban areas. Their strong financial foundation and consistent household routines make them well-suited for ongoing giving. Additionally, these groups contribute a significant portion of the total funds raised, particularly Groups C and I (fig.19).
- Groups A (First Class Life), F (Establishing Roots), and D (Secure Tranquillity) follow closely, characterized by financial stability and professional occupation (fig.18). These segments also have a moderately high number of supporters and contribute substantial donation amounts (fig.19)
- Other groups such as H (Middle Blue-collars), K (Mature Freedom), and L (Rural Commitment) have relatively disadvantaged financial foundations and represent the lowest number of total supporters. However, they exhibit a notably high proportion of regular givers. This suggests that while the overall supporter base in these segments is small, those who do give are likely to have stable income and a strong commitment to long-term giving. In contrast, the majority in these groups may not engage at

all, which results in low total supporter numbers but a high conversion rate among the engaged minority.

- Segments with the largest number of regular givers (H, L and B) often show average or below-average conversion rates, while smaller segments regularly achieve disproportionately high conversion (D and I). Additionally, this is a consistent trend observed across all groups.

V. Drawing business links to potential relationships

1. Low-value, high-frequency donors are key to sustainable revenue

- **Observed relationship:** Regular givers typically contribute smaller amounts per transaction but cumulatively deliver higher total revenue due to consistent giving. Furthermore, RFM analysis suggests that frequent and smaller donations correlate with higher regular givers proportion compared to infrequent, larger donations, indicating that monetary value alone does not strongly predict regular giving status.
- **Business Insights:**
  - The value of regular donors arises primarily from **sustained engagement across many donors** rather than depending on a limited number of high-value contributors. This highlights an opportunity to broaden donor segments by promoting **affordable, recurring donations**.
  - Focusing only on high-value donors may overlook the stability provided by a broader base of regular, lower-value contributors and increase dependence on costly acquisition drives to replace inconsistent one-off donations.
  - **Encouraging consistent giving behaviours** can enhance revenue predictability and supports long-term program funding.

2. Frequency is a strong indicator of conversion potential

- **Observed relationship**: The proportion of regular givers increases steadily with higher frequency scores, indicating that **how often** donors give is a stronger predictor of regular giving than **how much** they give each time.
- **Business Insights:** Targeting RFM groups with an already high proportion of regular givers is likely to improve outreach efficiency. Donors within the same group tend to exhibit similar behaviors, suggesting that non-regular givers in these segments **closely resemble regular givers.** As a result, they may be **more responsive to follow-up communication if approached strategically**.

3. Mid-to-high RFM scores signal high conversion potential

- **Observed relationship**: Higher combined RFM scores (particularly from mid-range scores of around 7 upward) show increasingly higher regular giving and conversion rates.
- **Business Insights:** Donors with **mid-to-high RFM** scores represent segments with **high conversion potential**, pointing toward opportunities for **targeted campaign investments** to improve overall fundraising efficiency and return on investment.

4. Tailored Strategies Needed for Different MOSAIC Groups and Regions

- **Observed relationship**: Certain regions, such as South Australia and Tasmania, exhibit higher conversion rates despite smaller donor bases. Additionally, MOSAIC groups show varying behaviours with some having high regular giver rates but low conversion rates, and vice versa.
- **Business Insights:**
    - The highest total contributions come from both financially stable groups (A and C) and a group typically considered disadvantaged (N). This pattern may be explained by the possibility that donors in Group N represent a wealthier minority within an otherwise disadvantaged segment, supported by the relatively small number of supporters in group N.
    - A **one-size-fits-all approach** is not optimal. UNICEF should develop **tailored acquisition and retention strategies** that consider the unique characteristics of each MOSAIC group or region.
    - High regular giver regions may benefit from campaigns aimed at increasing donor headcount, while regions with low conversion rates might require strategies focused on improving conversion efficiency.

5. Notably high product loyalty and low donor migration

- **Observed relationship:** Regular donors often remain loyal to the specific program or product they initially supported, particularly the **Global Parent program.**
- **Business Insights:**
    - This program can be positioned as **a key acquisition channel** for converting one-time givers into regular givers. Efforts should go into optimizing and promoting this program as a starting point for sustained engagement.
    - However, Global Parentis primarily designed for **donor acquisition** rather than the conversion of existing supporters. This trend may suggest that most conversions are self-driven by donors' motivations rather than through active conversion strategies, indicating that **conversion efforts are limited.**

VI. Executive Briefing

1. Key findings

- **Frequent, low-value donors Provide Financial Stability**
  - Donors who contribute small amounts consistently (e.g., \$10–\$30 monthly) tend to generate more cumulative value over time than those who make larger, one-off donations.
  - These supporters exhibit frequent engagement and predictable behaviour, collectively forming **a financially resilient donor base.**
- **Frequency of giving is strongly related to conversion potential**
  - Donors with moderate to high recent activity show strong potential for conversion to regular giving.
  - Their behavioural patterns closely resemble existing regular givers, suggesting they are more likely to convert if approached with the right engagement strategy.
- **Conversion outcomes differ by region and demographic group**
  - South Australia and Tasmania demonstrate above-average donor conversion rates despite smaller bases, potentially implying localized strategies are particularly effective.
  - Certain demographic segments, including suburban families and older professionals, tend to engage well once converted.
- **The Global Parent program attracts and retains the majority of regular givers**
  - Global Parent, a recurring family-focused campaign, shows high donor retention. However, their conversions appear to be driven by donors' self-motivation rather than by conversion efforts, indicating room for improvements in proactive acquisition strategy.

2. Suggestions

- **Encourage and maintain frequent engagement:** This can be achieved by developing **frequency-based targeting models** such as reaching donors with 3+ donations in the last year. Messaging should emphasize how consistent small giving can create big impacts. This approach helps strengthen the regular giver base, ensuring **more stable income** while **reducing reliance on one-off donors.**
- **Customize messages and campaign visuals to target specific regions or demographic groups:** Tailor acquisition campaigns to attract new supporters in high-conversion regions, where donors are more likely to commit to regular giving. Simultaneously, focus retention efforts on existing donors in low-conversion areas. This targeted approach leverages available data to make informed decisions, improve outreach efficiency, and lower overall fundraising costs.
- **Increase donor retention campaigns with clear targeted individuals:** Identify one-time donors with 2–3 recent donations and show similar behavioural patterns to regular givers to be the main target.

3. Potential further investigations

- **Understanding donor insights:** Conduct targeted surveys in high-performing regions and demographic groups to uncover motivations behind recurring giving and factors that contribute to donor loyalty. This will help improve future campaigns and shape more resonant messaging.

Appendix

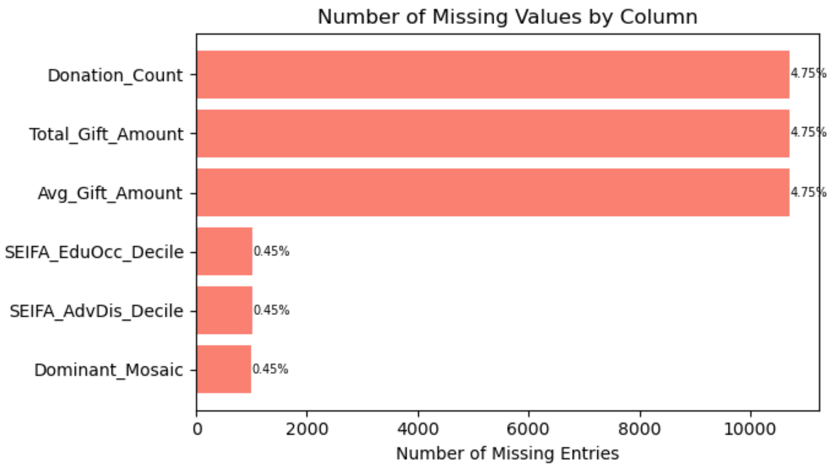
Appendix A: Original datasets summary

Dataset	Rows	Columns
Supporter	661292	18
Campaigns	21943	10
Donations	6114649	8
Regular Giving	158301	14

Appendix B: External dataset, ABS SEIFA

Data link: https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release#data-downloads
Data name: *Postal Area, Indexes, SEIFA 2021.xlsx*

Appendix C: Missing values of final dataset



Appendix D: Chi-Squared test contingency tables

Table 1: Observed counts

Mosaic_Group	Non-regular	Regular
A	14660	7023
B	7939	5123
C	17806	9464
D	13147	5630
E	5684	2758
F	13505	7912
G	16940	10684
H	3218	2874
I	16468	8510
J	7072	4575
K	9712	4903
L	3142	2550
M	5186	2666
N	2758	1696

Table 2: Expected counts

Mosaic_Group	Non-regular	Regular
A	13930.90	7752.10
B	8392.08	4669.92
C	17520.44	9749.56
D	12063.85	6713.15
E	5423.82	3018.18
F	13760.00	7657.00
G	17747.88	9876.12
H	3913.99	2178.01
I	16047.87	8930.13
J	7482.97	4164.03
K	9389.85	5225.15
L	3657.00	2035.00
M	5044.76	2807.24
N	2861.61	1592.39

References

Australian Bureau of Statistics. (2023, September 5). *Socio-Economic Indexes for Areas (SEIFA)*. Www.abs.gov.au. https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release#index-of-relative-socio-economic-advantage-and-disadvantage-irsad-

Australian Charities and Not-for-profits Commission. (2021, September 1). *Australian Charities and Not-for-profits Commission Act 2012*. Www.legislation.gov.au; scheme=AGLSTERMS.AglsAgent; corporateName=Office Parliamentary Counsel; address=Locked Bag 30 Kingston ACT 2604; contact=+61 2 6120 1400. https://www.legislation.gov.au/C2012A00168/latest/text

Bagot, K. L., Murray, A. L., & Masser, B. M. (2016). How Can We Improve Retention of the First-Time Donor? A Systematic Review of the Current Evidence. *Transfusion Medicine Reviews*, *30*(2), 81–91. https://doi.org/10.1016/j.tmrv.2016.02.002

Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM Ranking – an Effective Approach to Customer Segmentation. *Journal of King Saud University - Computer and Information Sciences*, *33*(10). https://doi.org/10.1016/j.jksuci.2018.09.004

Kuhn, M., & Johnson, K. (2013). Data pre-processing. In *Applied predictive modeling* (pp. 27–59). Springer New York. https://doi.org/10.1007/978-1-4614-6849-3

Reuters. (2025, April). UNICEF projects 20% drop in 2026 funding after US cuts. *Reuters*. https://www.reuters.com/world/unicef-projects-20-drop-2026-funding-after-us-cuts-2025-04-15/

Singh, A. S., & Masuku, M. (2014). Sampling techniques and determination of sample size in applied statistics research: An overview. *International Journal of Commerce and Management*, *2*, 1–22.