BA Project Report :Part A (Individual)

Name: Ayman Ahmed

Student ID: 104497350

INF70008-Business Analytics and Visualisation

Assignment 1 - BA Project Report - Part A (Individual)

Due Date: 9th April 2024

Assignment Value = 35%

Table of Contents

1 Introduction

This report dives into an assessment of the data quality and analytics concerning donation trends for Reachout Worldwide (ROW). The analysis covers various aspects including donor segmentation, retention rates, churn rates, preferred donation channels, donor activity throughout the year, and appeal success. Through insightful visualizations and key insights, this report aims to provide actionable recommendations for ROW to optimize donation strategies and enhance fundraising efforts.

2 Data Quality Assessment

The term “data wrangling” emerged in the late 1990s and aptly describes the steps data teams take to prepare data for model development.

Similar to rounding up and organizing livestock, data wrangling involves gathering diverse data sources and data cleaning corrects the inconsistencies (Rosett & Hagerty, 2021).

The dataset in general was not heavily skewed and hence, it did not need much wrangling and cleaning. However, the impact of some of the cleaning tasks cannot be understated. Some of the appeals with large number of donations were not named properly such as app_002. Additionally, many of the countries were inconsistently named and hence corrected. Lastly, the duplicates removed were 4 in total and was rectified by excluding one of them for each instance.

Sr No	Data Quality Issues	Field	Data Affected	Wrangling Solution
1	Inconsistent naming – ‘Singapore’ was named as ‘Singpore’	Country of Residence (Donors Table)	Donor IDs VAL-5029, VAL-5464 (approx.. <1% of the data)	Group both values together to one value named ‘Singapore’
2	Inconsistent naming – ‘United States of America’ was named as ‘USA’	Country of Residence (Donors Table)	Donor IDs VAL-1014, VAL-12603, etc. (approx.. <1% of the data)	Group both values together to one value named 'United States of America’

3	Inconsistent naming – ‘China’ was named as ‘Republic of China’	Country of Residence (Donors Table)	Donor IDs VAL-10104, VAL-1036. etc. (approx.. <1% of the data)	Group both values together to one value named 'China’
4	Spelling Error – ‘Viet Nam’	Country of Residence (Donors Table)	Donor IDs VAL-14783, VAL-2699, etc. (approx.. <1% of the data)	Corrected spelling to ‘Vietnam’
5	Inconsistent naming of Appeal Codes (Codes were named ‘0App_002’, ‘0App-004’, ‘0App_009’)	Appeal Code (Donations Data Extract Table)	Donation IDs DID-23-12628, DID-23-12807, etc. (approx.<1% of the data)	Corrected spelling and grouped the codes to App_002, App_004 and App_009
6	Inconsistent naming of Donor ID VAL-16748 (was named as ‘Valued Donor 16748’)	Donor ID (Donations Data Extract Table)	Donation ID DID-21-06906 (approx.<1% of the data)	Corrected spelling and grouped with VAL-16748
7	Duplicate row with ID DID-20-02676 removed because all its values match with DID-20-02670	Donations Data Extract Table	DID-20-02676 (only one row)	Excluded one entire row with ID DID-20-02676 as all of its fields except the ID matched with row with ID DID-20-02670,

				which proves that one of them is a duplicate row.
8	Duplicate row with ID DID-21-06354 removed because all its values match with DID-21-06243	Donations Data Extract Table	DID-21-06354 (only one row)	Excluded one entire row with ID DID-21-06354 as all of its fields except the ID matched with row with ID DID-21-06243, which proves that one of them is a duplicate row
9	Duplicate row with ID DID-22-03584 removed because all its values match with DID-22-03534	Donations Data Extract Table	DID-22-03584 (only one row)	Excluded one entire row with ID DID-22-03584 as all of its fields except the ID matched with row with ID DID-22-03534, which proves that one of them is a duplicate row

10	Duplicate row with ID DID-22-02309 removed because except **payment type**, all of its values match with DID-20-04460	Donations Data Extract Table	DID-22-02309 (only one row)	Excluded one entire row with ID DID-22-02309 as all of its fields except the ID and **payment type** matched with row with ID DID-20-04460. In this case, the system might have made an error, or it could have been a data entry mistake, and hence the same record was recorded twice, perhaps to correct the record.

3 Analytics & Visualizations

Across social science research, visualizations have been shown to improve comprehension and decision-making in numerous studies (Eberhard, 2021).

This part will explore some of the key insights we can infer from the sample dataset provided by ROW, by analysing and subsequently visualizing the data to illustrate the insights.

Key Insight 1: Donor Segmentation by Donation Amount & Frequency

The below visualization segments donors into four categories, based on their donation amount and their frequency. The four categories are:

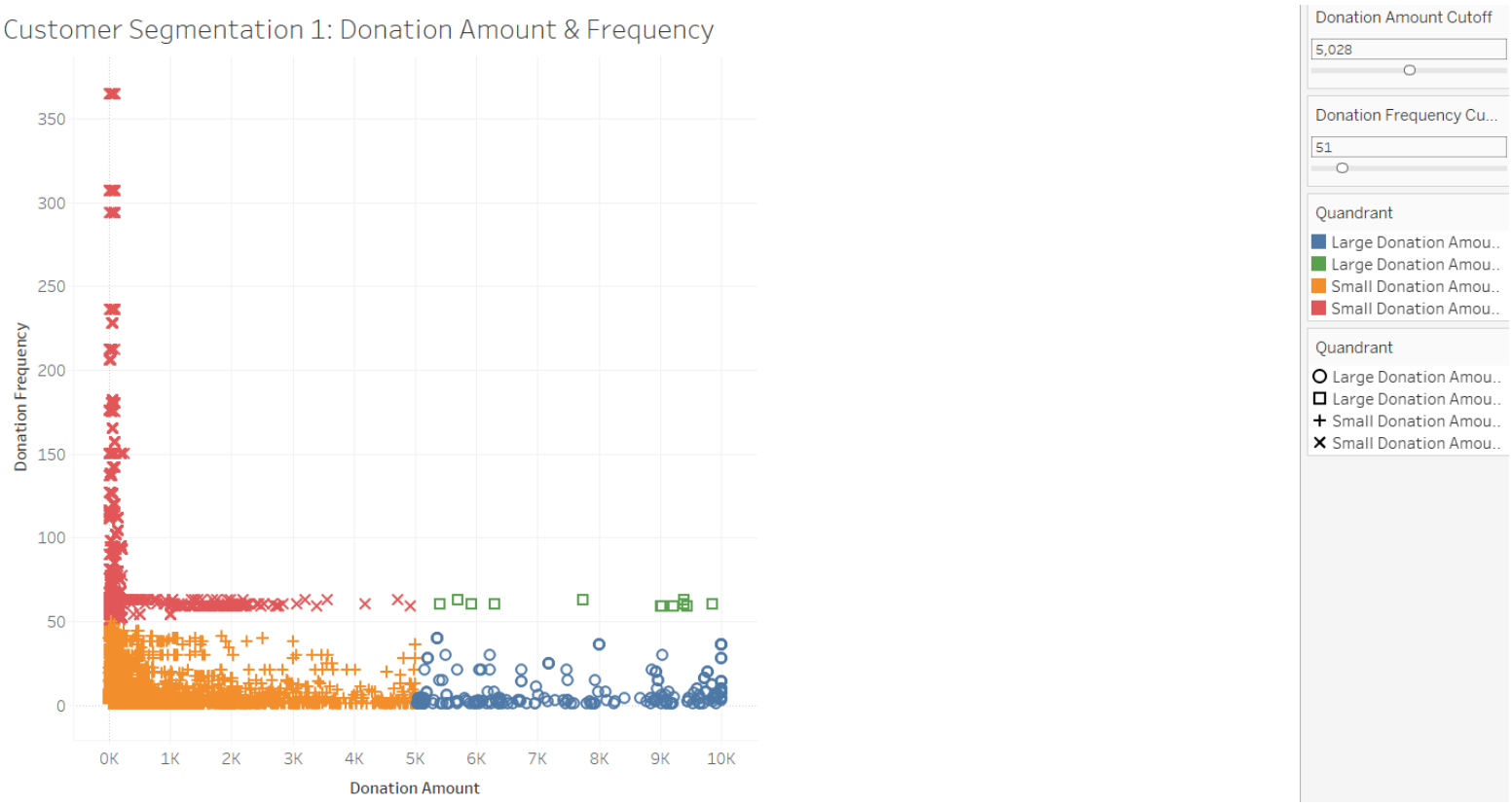
1)Large donation amount, less frequency (denoted by blue circles)

2) Large donation amount, more frequency (denoted by green squares)

3) Small donation amount, more frequency (denoted by red X's)

4) Small donation amount, less frequency (denoted by orange plus signs)

The donation amount cutoff was taken as the rough median, which is 5028, while the donation frequency cutoff was taken as 51. These cutoffs decided whether a donor was large/small and more/less frequent.

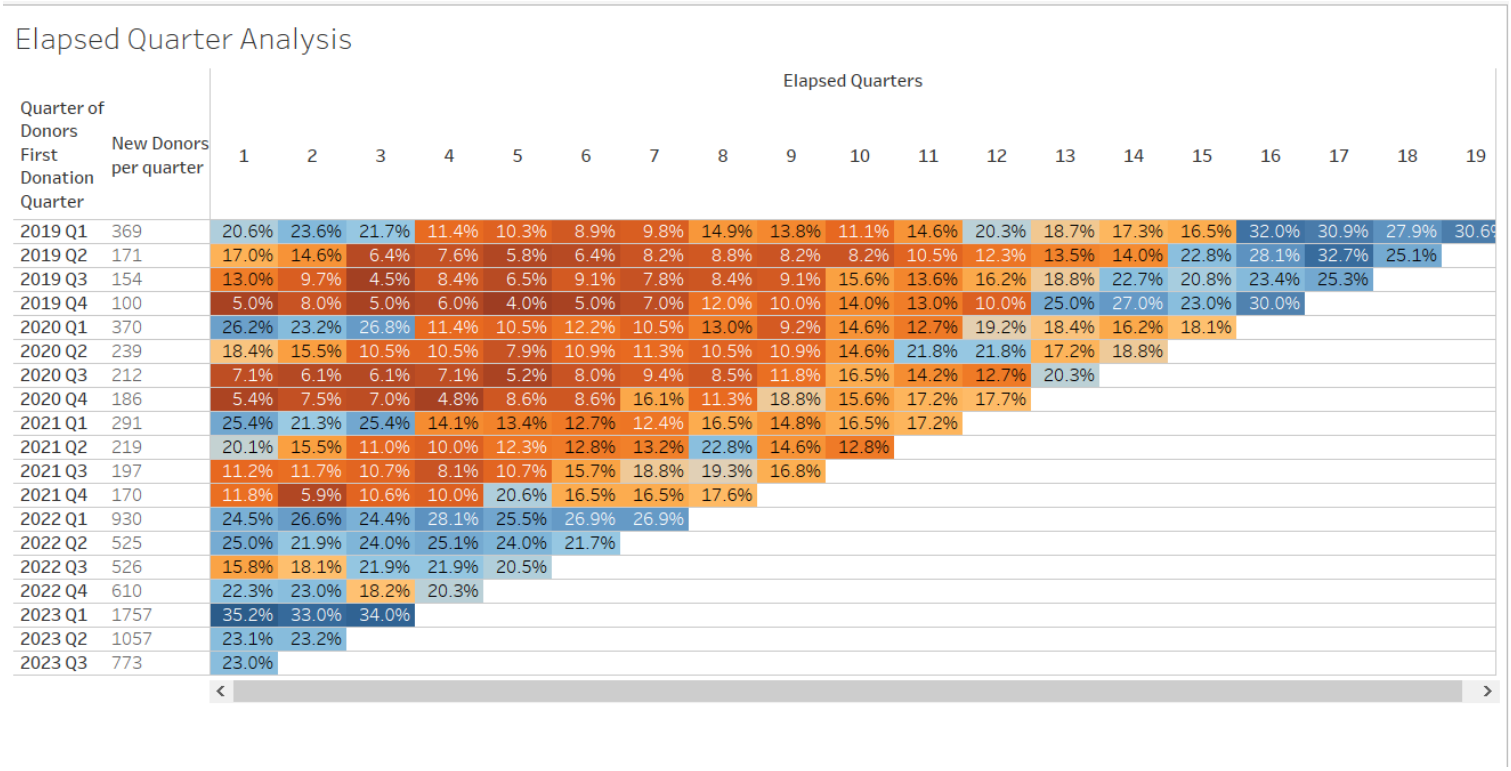


The quadrant shows that the Reachout Worldwide are receiving most of their donations in the small donations category. To drill down further small donations are being given with less frequency for the most part. On the other hand, large donations with high frequency are quite low.

Hence, we can infer that ROW should shift their focus to individuals and perhaps organizations with high donating capability to maximize their donations. Additionally, ROW should also try to retain their small donors to increase frequency of small donations.

Key Insight 2: Retention Rate Analysis

The below visualization shows the retention rates as quarters in the 2019-2023 period pass by.



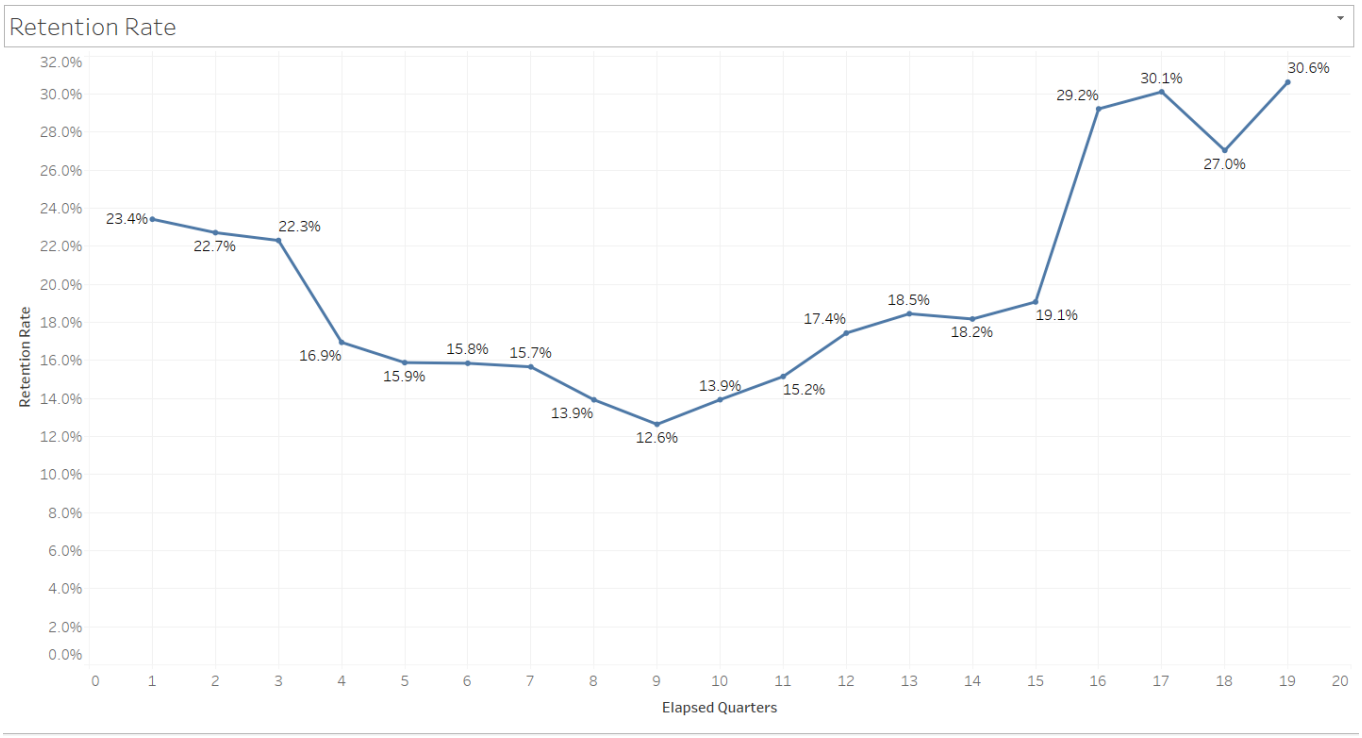
Elapsed Quarters

Quarter of Donors First Donation Quarter	New Donors per quarter	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2019 Q1	369	20.6%	23.6%	21.7%	11.4%	10.3%	8.9%	9.8%	14.9%	13.8%	11.1%	14.6%	20.3%	18.7%	17.3%	16.5%	32.0%	30.9%	27.9%	30.6%
2019 Q2	171	17.0%	14.6%	6.4%	7.6%	5.8%	6.4%	8.2%	8.8%	8.2%	8.2%	10.5%	12.3%	13.5%	14.0%	22.8%	28.1%	32.7%	25.1%	
2019 Q3	154	13.0%	9.7%	4.5%	8.4%	6.5%	9.1%	7.8%	8.4%	9.1%	15.6%	13.6%	16.2%	18.8%	22.7%	20.8%	23.4%	25.3%		
2019 Q4	100	5.0%	8.0%	5.0%	6.0%	4.0%	5.0%	7.0%	12.0%	10.0%	14.0%	13.0%	10.0%	25.0%	27.0%	23.0%	30.0%			
2020 Q1	370	26.2%	23.2%	26.8%	11.4%	10.5%	12.2%	10.5%	13.0%	9.2%	14.6%	12.7%	19.2%	18.4%	16.2%	18.1%				
2020 Q2	239	18.4%	15.5%	10.5%	10.5%	7.9%	10.9%	11.3%	10.5%	10.9%	14.6%	21.8%	21.8%	17.2%	18.8%					
2020 Q3	212	7.1%	6.1%	6.1%	7.1%	5.2%	8.0%	9.4%	8.5%	11.8%	16.5%	14.2%	12.7%	20.3%						
2020 Q4	186	5.4%	7.5%	7.0%	4.8%	8.6%	8.6%	16.1%	11.3%	18.8%	15.6%	17.2%	17.7%							
2021 Q1	291	25.4%	21.3%	25.4%	14.1%	13.4%	12.7%	12.4%	16.5%	14.8%	16.5%	17.2%								
2021 Q2	219	20.1%	15.5%	11.0%	10.0%	12.3%	12.8%	13.2%	22.8%	14.6%	12.8%									
2021 Q3	197	11.2%	11.7%	10.7%	8.1%	10.7%	15.7%	18.8%	19.3%	16.8%										
2021 Q4	170	11.8%	5.9%	10.6%	10.0%	20.6%	16.5%	16.5%	17.6%											
2022 Q1	930	24.5%	26.6%	24.4%	28.1%	25.5%	26.9%	26.9%												
2022 Q2	525	25.0%	21.9%	24.0%	25.1%	24.0%	21.7%													
2022 Q3	526	15.8%	18.1%	21.9%	21.9%	20.5%														
2022 Q4	610	22.3%	23.0%	18.2%	20.3%															
2023 Q1	1757	35.2%	33.0%	34.0%																
2023 Q2	1057	23.1%	23.2%																	
2023 Q3	773	23.0%																		

The first thing we can notice is that retention rates of donors are low after 3-4 quarters pass by since their first donation. However, the rates increase again steadily after about 10-11 quarters.

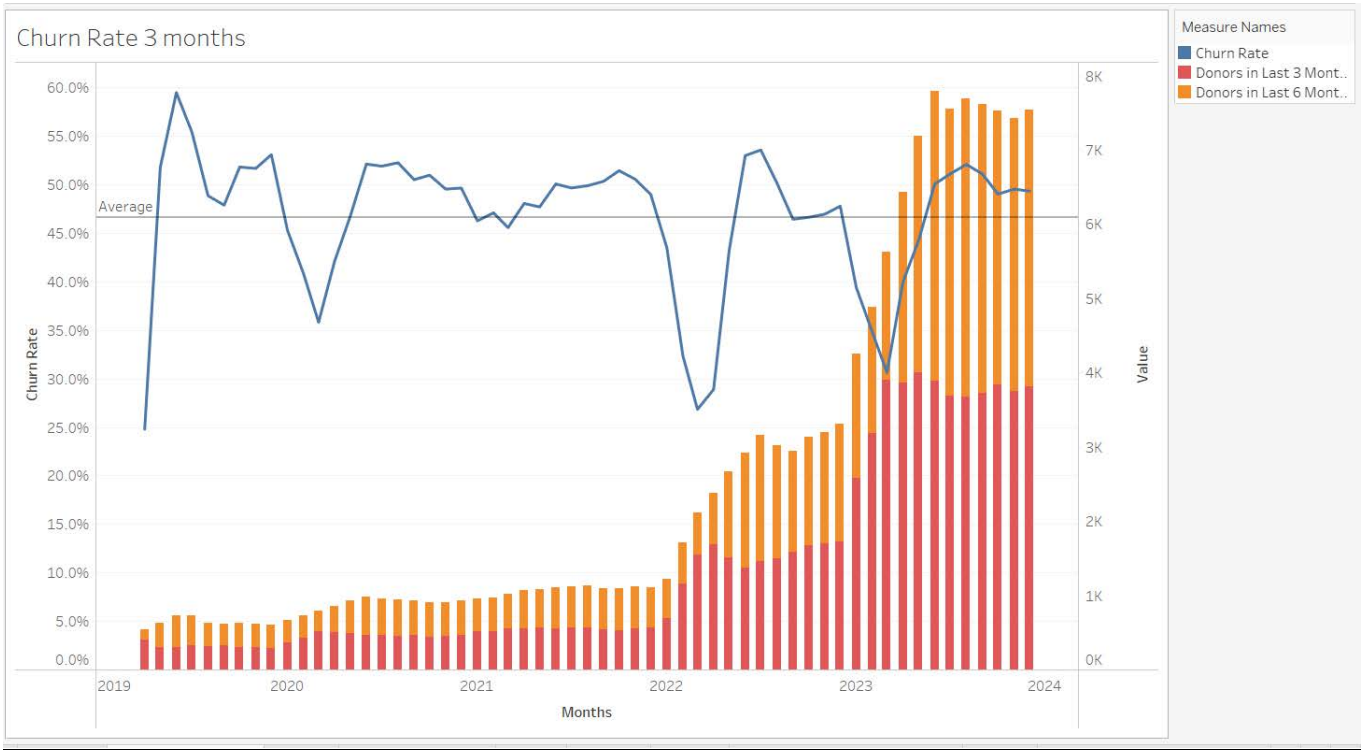
Another aspect that can be pointed out is that retention rates have increased 2022 onwards. Additionally, there is also a sharp increase in new donors since the start of 2022. Conclusively, there is an upward trend for number of new donors and the retention rates since 2022. ROW must increase their efforts to retain the surge in the new donors to capitalize on this opportunity.

The next visualization also supports the above.


Retention Rate
Retention Rate
Elapsed Quarters
32.0%
30.0%
28.0%
26.0%
24.0%
22.0%
20.0%
18.0%
16.0%
14.0%
12.0%
10.0%
8.0%
6.0%
4.0%
2.0%
0.0%
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
23.4%
22.7%
22.3%
16.9%
15.9%
15.8%
15.7%
13.9%
12.6%
13.9%
15.2%
17.4%
18.5%
18.2%
19.1%
29.2%
30.1%
27.0%
30.6%

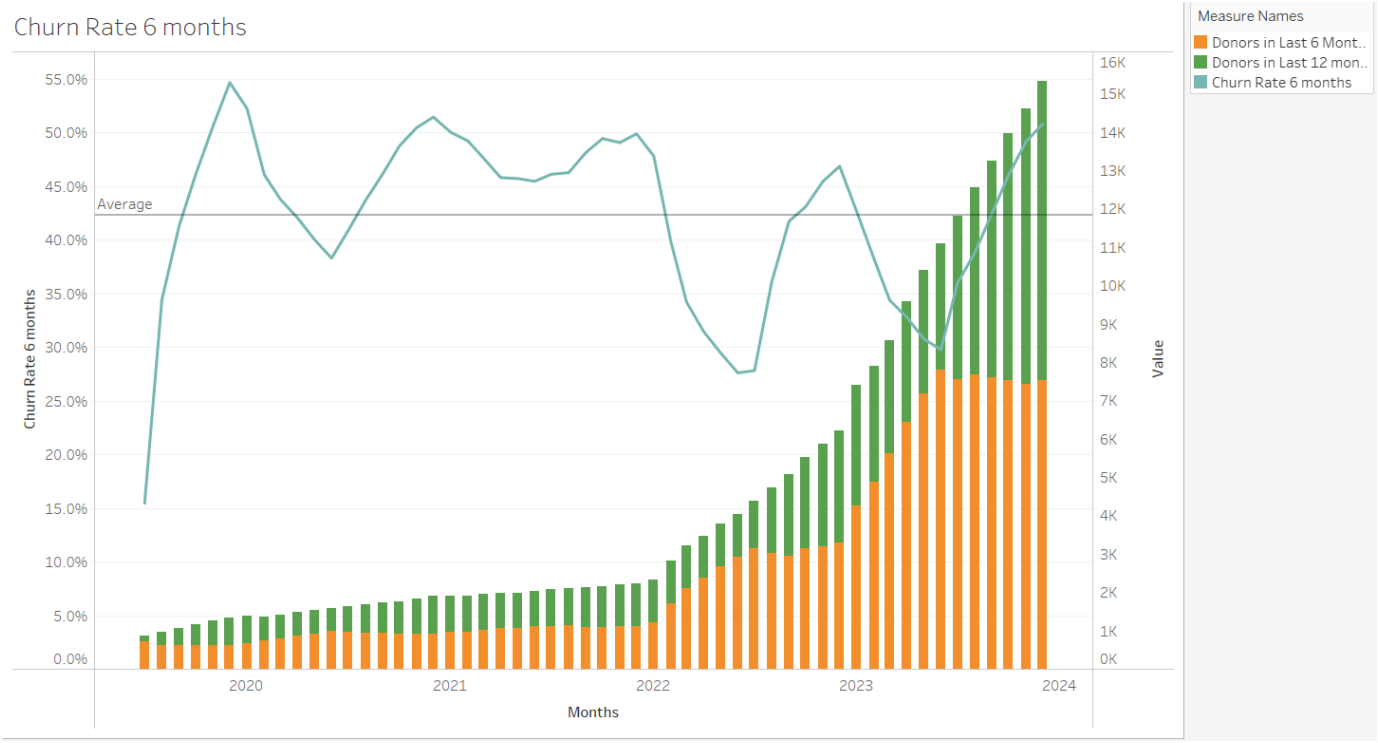
Key Insight 3: Donor Churn Rates

This section will illustrate the churn rate of the donors as the years pass by. First, we have a visualization showing the churn rate for every 3-month period. For added context, the red vertical stacks show the total number of donors in the last 3 months, while the orange stacks show the total number of donors in the last 6 months. Additionally, the average line shows the average churn rate (46.7%).



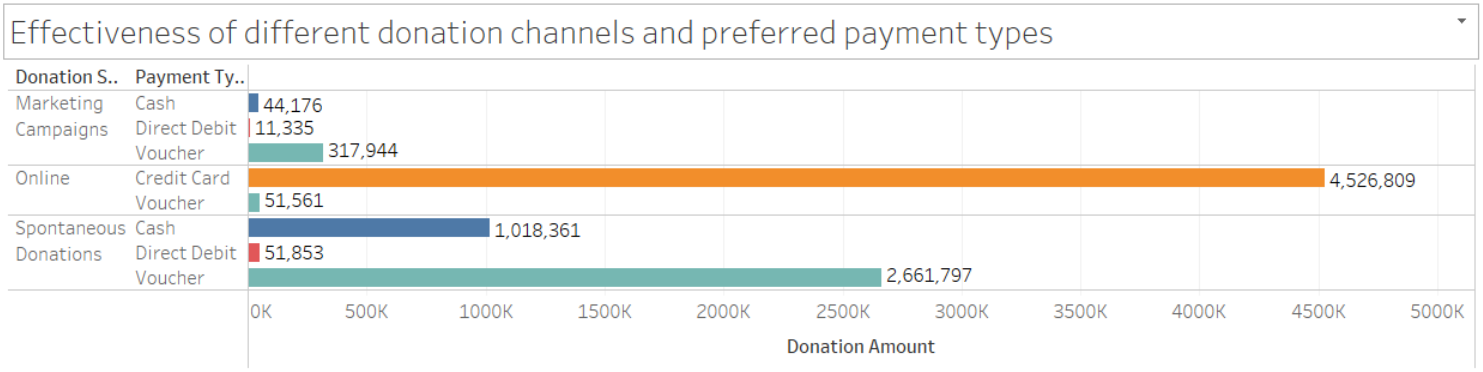
One of the most significant discoveries here is that the churn rate continually increases as the summer period approaches, the peaks being in June of every year to be exact, reaching up to almost 60%. ROW must focus on maximizing their efforts during this time, to decrease the number of donors lost during this period.

Similarly, in the 6 month churn rate visualization shown below, there is recurring trend of the churn rate increasing every winter, the peaks being in December of each year to be exact and reaching almost 55% in December 2019.



Key Insight 4: Effectiveness of different donation channels and preferred payment types

This visualization will reveal the preferred payment types across different donation channels.



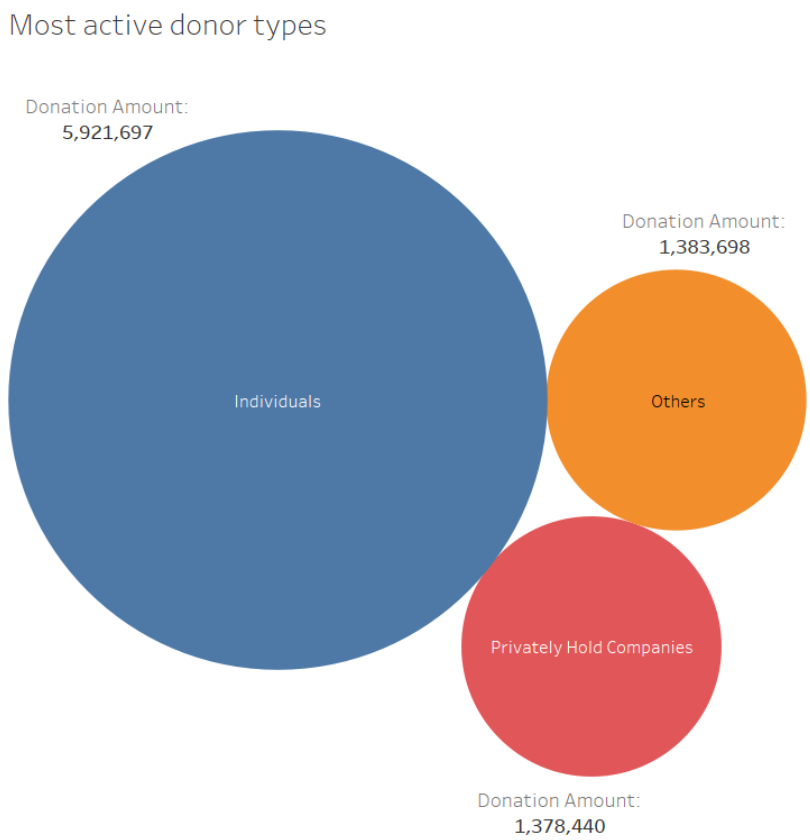
At a glance, online transactions seem to be extremely popular for donations, and hence ROW should continue to advertise and perhaps increase advertising online to raise more funds.

Marketing campaigns don't seem to be popular at all in terms of total donations received. However, vouchers within the campaigns seem to be popular.

Additionally, spontaneous donations are also largely received through vouchers. Hence promoting vouchers might be useful in campaigns and to encourage more spontaneous donations.

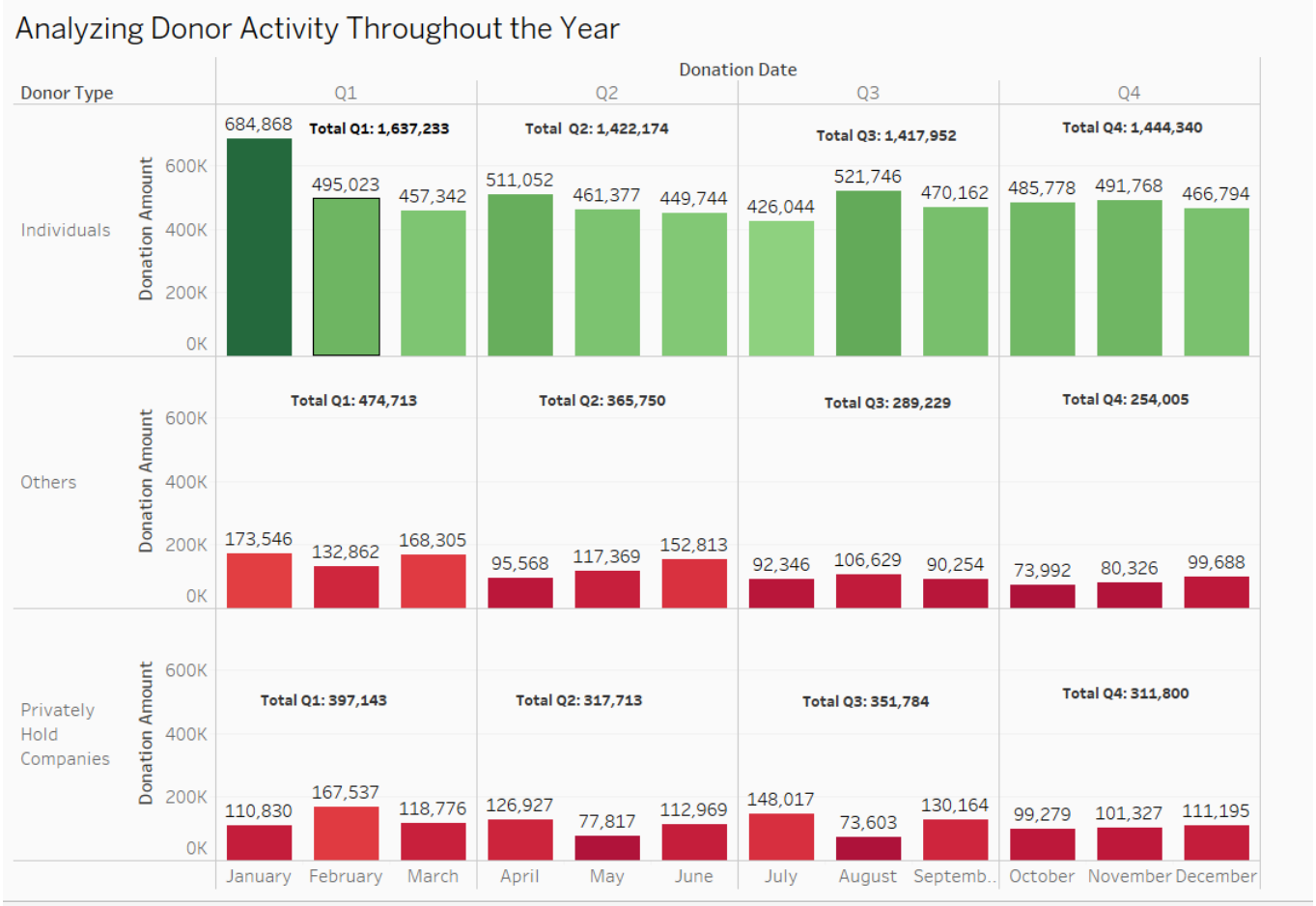
On the other hand, people do not prefer using vouchers online, likely due to the ease of credit card payments as proven by the large amount of donations received through credit card.

Key Insight 5: Analyzing Donor Activity Throughout the Year



Clearly, individuals are the most active donors with almost 6 million total donations across 4 years. The second most and third most active donors are close, with only 5000 AUD separating them.

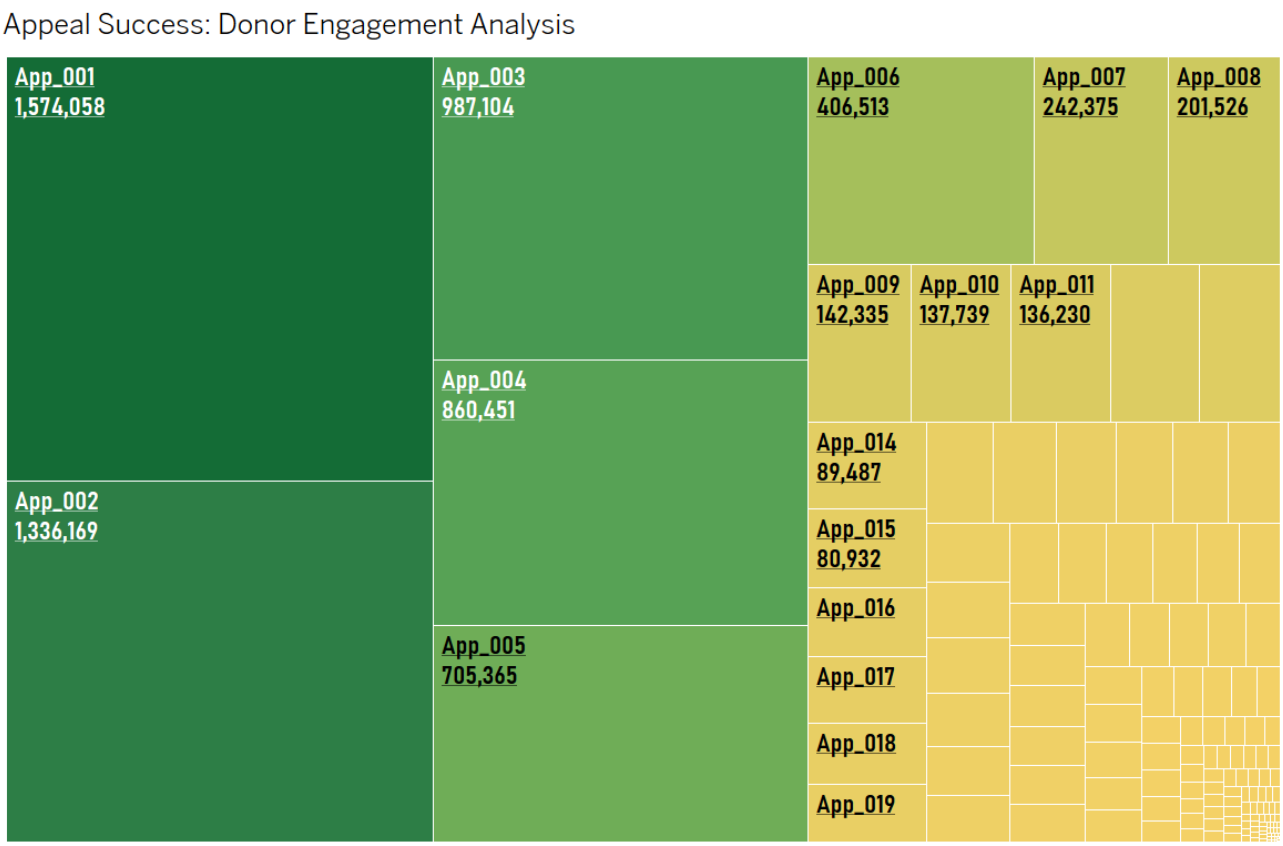
A further drill down can analyse their peak activity in different times of the year:



A trend within this visualization is that during the first quarter of every year, ROW receives the most donations as compared to the rest of the year. This trend is consistent across all donor types.

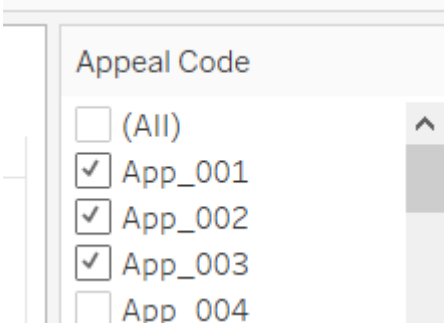
Considering this information, ROW must increase their campaigns and advertising during the rest of the year, especially during the end of the year, where companies and others donate the least as compared to the rest of the year.

Key Insight 6: Appeal Success Analysis



The above visualization clearly shows appeals 001, 002 and 003 were the top 3 most popular appeals.

Next, we drill down to see which appeals resonated with which donor types across different geographical regions. The appeals were filtered out to only include the top 3 appeals.



Across all donor types, the top 3 appeals (001,002 and 003) were most popular in the USA. ROW can infer from this that appeals of this kind should be directed towards the USA, and more specifically, they should target individuals and private companies since they were the donor types with most donations.

3 Conclusion

In conclusion, the report highlights the importance of data-driven decision-making in enhancing ROW's fundraising endeavors. By leveraging insights from donor segmentation, retention rates, churn rates, preferred payment types, donor activity, and appeal success, ROW can refine its strategies to maximize donation outcomes. Implementing targeted campaigns, focusing on high-value donors, and optimizing donation channels are key recommendations to sustain and enhance fundraising efforts. Through continuous analysis and adaptation, ROW can effectively engage donors and drive impactful contributions towards its mission.

References

Eberhard, K. (2021). The Effects of Visualization on Judgment and decision-making: a Systematic Literature Review. *Management Review Quarterly*, *73*. https://doi.org/10.1007/s11301-021-00235-8

Rosett, C. M., & Hagerty, A. (2021). Data Wrangling. *Springer EBooks*, 217–241. https://doi.org/10.1007/978-3-030-67626-1_13