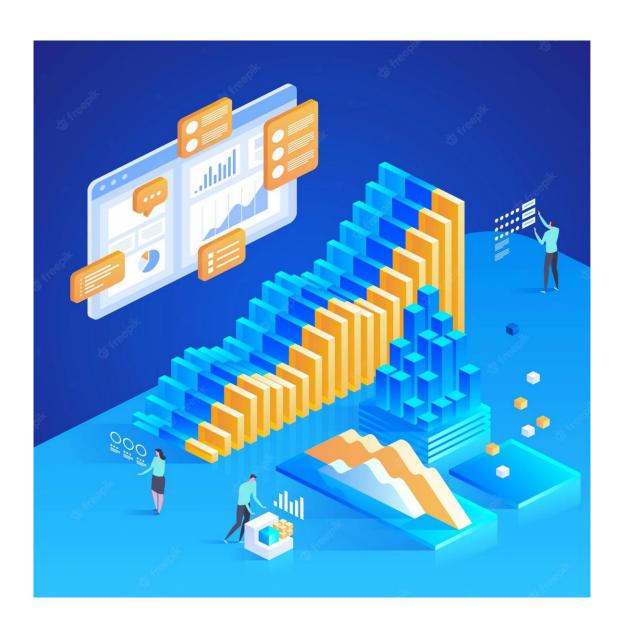
# Education For All: Donor and Donation Analytics

EntryLevel Data Analyst 2:SQL Project



## **Professional Background**

My name is Ayokunle James, a graduate of Mechanical engineering, and a data /business intelligence analyst.

Data and data analytics piqued my interest while conducting research for my final undergraduate thesis where I recorded thousands of data points and performed simple analysis of the datasets.

After graduation in 2020, I worked as an operations manager at a production facility where I analyzed and visualized production & operations data using spreadsheets (Google Sheets).

I am currently a Business Intelligence Analyst at an e-commerce startup. In this role, I generate reports and create dashboards in Metabase by writing SQL queries. I liaise with business stakeholders to understand their individual problems and requirements, then build solutions to these problems.

My skills include data gathering, data cleaning, data exploration, data analysis and data visualization. Tools include Excel, SQL (PostgreSQL, MySQL, SQLite), PowerBI, Metabase. I am looking for more opportunities to learn and develop myself in the field of data analytics.

## Table of Contents

Professional Background	
Introduction	3
Situation	3
Task	3
Action	3
Result	5
Root Cause Analysis	6
Concept	6
5 Why's Process	6
Insights	9
Donor and Donation Value Insights	9
Demographic Insights	12
Additional Insights	15
Visualizations	18
Findings & Recommendations	20
Findings	20
Recommendations	21
Conclusion	22

#### Introduction

#### Situation:

I am a Data Analyst working for the charity, *Education for All*. I have been asked by the Head of Fundraising to present the data on donor insights and donation rates. *Education for All* relies solely on funds from well-meaning individuals in form of monetary donations. These donations fund the activities and missions of the organization.

In the year in review, the organization has received donations from individuals across different genders, job fields, languages, educational status, and locations in the USA. In two (2) weeks, the fundraising team is having a fundraising strategy meeting to achieve certain objectives in the following year. The objectives are to:

- Increase the number of donors in the database.
- Increase the donation frequency of donors.
- Increase the value of donations in your database.

#### Task:

The aim of this project is to provide insights from the available donation data to inform the fundraising strategy and increase donations for the following year. I had to properly frame and define the business problem, the data that needed to be collected to understand the problem, and the questions that needed to be asked to understand the problem.

#### Action:

The data provided was relatively clean. I checked for duplicate rows, blank rows, and misspellings, and 0 rows were returned. I also checked to ensure that each column had the appropriate data type. After properly framing the business problem, I dived into the data to extract insights relevant to the problem, which will help the fundraising team to properly strategize. I also did a Root Cause Analysis (RCA) of the business problem to dig deeper into the causes of the problem. I then created visualizations of the insights extracted for presentation to the fundraising team during the strategy meeting. The tools used for this project:

- Data Cleaning: Microsoft Excel
- Database Management System: SQLite Online
- Data Visualization: Microsoft PowerBI

The organization collects and stores data about donors in 2 datasets:

- Donation\_Data: fields include id (Donor ID), first\_name (Donor first name),
   last\_name (Donor last name), email (Donor email address), gender (Donor gender),
   job\_field (Donor job field), donation (Donation amount), state (Donor state of residence (US)),
   shirt\_size (Donor t-shirt size).
- Donor Data2: fields include id (Donor ID), donation\_frequency (Frequency of donation), university (Donor University attended), car (Donor car make), second\_language (Donor second language), favourite\_colour (Donor favourite colour), movie\_genre (Donor favourite movie genre).
- 3. I also created a new column **total\_donation\_value**, which computes the total value of donations of each donor in the year in review.

I imported both datasets into SQLite Online. I then joined both tables using an inner join, to have a broader insight on donors and their donations. For my source query, I only selected columns relevant to the analysis of the business problem. I ran the query and exported the result as a .csv file. All insights extracted from the database were drawn majorly from this query. The SQL query and resulting table are shown below.

```
1 SELECT
2 Donation_Data.id,
3 --first_name,
4 --last_name,
5 --email,
6 gender,
7 job_field,
donation,
9 donation_frequency,
10 case
11 when donation_frequency = 'Once' then (donation * 1)
12 when donation_frequency = 'Weekly' then (donation * 52)
13 when donation_frequency = 'Wenthly' Then (donation * 12)
14 when donation frequency = 'Yearly' Then (donation * 1)
15 Else 'Check Error'
16 End as total_donation_value,
17 state,
18 university
19 --car,
20 --second_language,
21 --favourite_colour,
22 --movie_genre,
3 --shirt_size
24 FROM Donation_Data
25 on Donation_Data
26 on Donation_Data
27 order by 1
```

Figure 1: Source Query

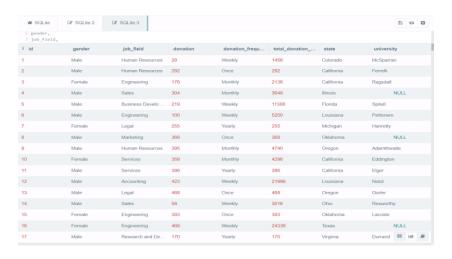


Figure 2: Source Query Result

# Result:

I observed trends in the data and extracted some interesting insights from the data on the donor, donation value, and donating demographics. These insights will help the fundraising team to properly strategize and increase donation value in the following year.

#### **Root Cause Analysis**

#### Concept

Root Cause Analysis is a systematic process of identifying the root causes of a problem and analyzing for possible ways to respond to them. This system works to nut just find a solution to a problem but also how to prevent the problem.

The 5 Whys technique is one of the most effective tools for root cause analysis in the Lean management arsenal invented by Sakichi Toyoda.

"The basis of Toyota's scientific approach is to ask why five times whenever we find a problem ... By repeating why five times, the nature of the problem as well as its solution becomes clear. "Taiichi Ohno

## 5 Why's Process

- 1. Why is *Education for All* having a fundraising strategy meeting to increase donations?
  - a. As a charity organization, *Education for All* relies on revenue from donations, and revenue in the year in review is low or insufficient to handle the charity's projects.
- 2. Why is the revenue low?
  - a. Revenue is low because the value of donations from donors in the year in review is low.
- 3. Why is the total value of donations low?
  - a. Small number of donors: There are 1000 donors in the NGO's database. This figure could be said to be low considering that the population of the USA is 332,403,650. Of the 1000 donors, a higher percentage (52.3%) either donated to the charity just once, or donate on a yearly basis, which is also once in the year in review.
  - b. Low frequency of donation: It can be observed from the donation data that only 24.5% of donors donate weekly, 23.2% donate monthly. 52.3% either donate yearly (25.9%) or donated just once (26.4%). The lower donation frequencies have the most donors. This means that most donors in the database donate less frequently.
  - c. Low value of donation: the data shows that value of donations from donors who donated once and donate yearly is low compared to those who donate monthly and weekly. The average donation value of donors is \$3922. The maximum value of donations from any donor who donated to the charity just once is \$491, and from donors who donated yearly (once in the year in review) is \$500. This is a problem because since the number of

donors who donate less frequently is higher than those who donate more frequently, the total donation value will be lower.

4. Why do we have small number of donors? Why is the frequency of donation low? Why is the value of donation low?

These problems could be due to a few mutual causes:

- a) High donor Churn: In business, customer churn analysis is a way to understand the number or percentage of customers who don't purchase additional products or services. In the context of a charity organization, churn may be defined as the number of donors who stopped donating to the charity. It is calculated by dividing the total from last year into the number of those same donors who haven't donated this year. When evaluated on an annual basis, the donor churn of *Education for All* in the new year may be estimated by the number of donors in the database who donated just once. These donors represent 26.4% of the total donors in the database. The high churn lowers the number of donors in the following year and consequently, the donation value.
- b) Low donor income: The income bracket of certain states and job fields is low, hence influencing the frequency and value of donations from such donors. Income bracket of a particular demographic may be deduced from the average value of donations from such demographic. 50.6% of donors work in job fields where the average donation is lower than the average total donation of donors in the database. 59.4% of donors live in states and work in job fields where the average donation is lower than the average total donation of donors.
- c) Small number of donors with low level of education: the level of education of donors could influence their donation frequency and donation amount to a charity like Education for All. Donors who didn't get a university education may tend to understand and resonate more with the impact of the work of Education for All and want to create better opportunities for others than they had themselves. From the data, I observed that the average donation of an uneducated donor is 12.9% higher than that of an educated donor. the higher % of uneducated donors also donated more frequently (weekly), than did the higher % of educated donors who only donated once.

#### 5. Why is there high donor churn?

- a) Inefficiency of the organization's charity projects: Donors would be unmotivated to donate to the organization if they realize that the quality and efficiency of the charity projects for which they are donating is low.
- b) Other Charity organizations: There are several other charity organizations competing for donations from the same donors.
- c) Inefficient outreach and follow-up efforts: 45 of the 50 states in the US have less than 40 donors per state, with only one state having up to 100 donors, while we have no record of donors in one state. It is noteworthy that of the top 10 states in terms of average donation value, only 3 rank in the top 10 in terms of total donation value. This may be due to population of donors, generosity of residents, and/or fundraising outreach and awareness. Regardless, this data reflects the low strength of the organization's outreach programs in each state.

## **Insights**

## Donor and Donation Value Insights

1. Donor and Donation Value: This query extracts insight on total number of donors, the minimum total donation value, the maximum total donation value, and average donation value.

```
SELECT

COUNT(3d) AS donor_count,

SUM(total_donation_value) AS total_donation,

round(AVG(total_donation_value).1) AS average_donation,

HING(total_donation_value) AS max_donation,

HING(total_donation_value) AS min_donation

FROM (SELECT

Donation_Data.id,

gender,

job_field,

donation,

donation_frequency,

CASE

WHEN donation_frequency = 'Once' THEN (donation * 1)

WHEN donation_frequency = 'Nece' THEN (donation * 52)

WHEN donation_frequency = 'Nece' THEN (donation * 12)

WHEN donation_frequency = 'Nece' THEN (donation * 12)

WHEN donation_frequency = 'Nenthly' THEN (donation * 12)

WHEN donation_frequency = 'Yearly' THEN (donation * 1)

ELSE 'Check Error'

END AS total_donation_value,

state

FROM Donation_Data

INMER JOIN Donor_Data2

ON Donation_Data.id = Donor_Data2.id )

i donor_count

total_donation

max_donation

max_donation

1000
```

Figure 3: Total Donor Count, Minimum Donation Value, Maximum Donation Value, and Average Donation Value.

2. Donation Frequency: This query aggregates number of donors, the minimum total donation value, the maximum total donation value, and average donation value by distinct donation frequency.

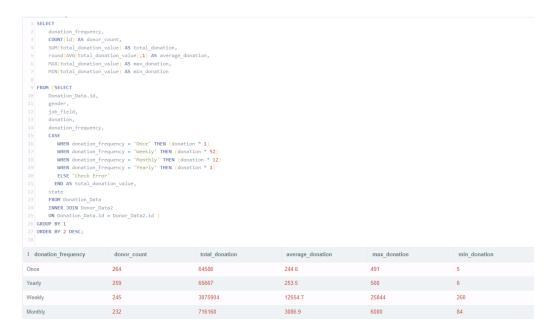


Figure 4: Donation Frequency data

3. Donor churn: This is the percentage of donors who stopped donating. This percentage may be estimated by the number of donors who donated to the charity just once, as opposed to committing. The query also extracts the percentage of total donation value that is lost to those donors.

```
1 SELECT
 2 COUNT(CASE WHEN donation frequency = 'Once' THEN id END)*1.0
 4 COUNT(DISTINCT id)*1.0 AS donor_churn,
6 round(SUM(CASE WHEN donation_frequency = 'Once' THEN total_donation_value END) *1.0
 8 sum( total_donation_value)*1.0,3) AS donation_churn
9 FROM(SELECT
      CASE
              WHEN donation_frequency = 'Once' THEN (donation * 1)
             WHEN donation_frequency = 'Weekly' THEN (donation * 52)
WHEN donation_frequency = 'Monthly' THEN (donation * 12)
             WHEN donation_frequency = 'Yearly' THEN (donation * 1)
             ELSE 'Check Error'
          END AS total_donation_value
      FROM Donation_Data
      INNER JOIN Donor_Data2
       ON Donation_Data.id = Donor_Data2.id
i donor_churn
                                                                            donation_churn
0.264
                                                                           0.016
```

Figure 5: Churn

## Demographic Insights

1. Gender distribution: Donor Count, Total Donation Value, Average Donation Value, Minimum and Maximum Donation Value.

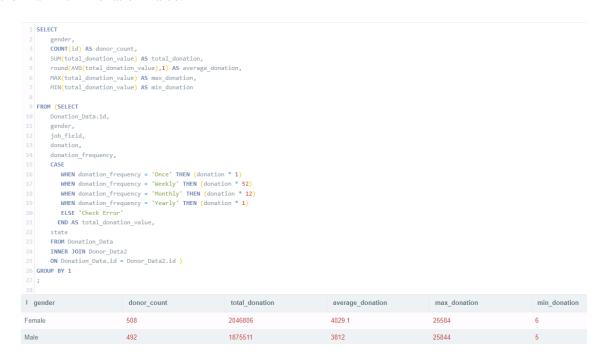


Figure 6: Donor Count, Minimum Donation Value, Maximum Donation Value, and Average Donation

Value by Gender

2. Least Donating States: this query extracts the states in ascending order of number of donors.

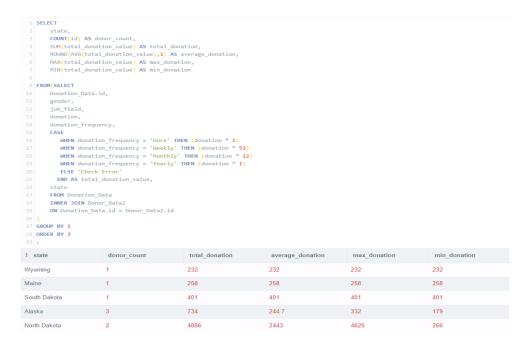


Figure 7: Lowest States by Total donation value

3. Top Job Fields by Average Donation Value: This query extracts job fields in order of descending donor count and average donation value.

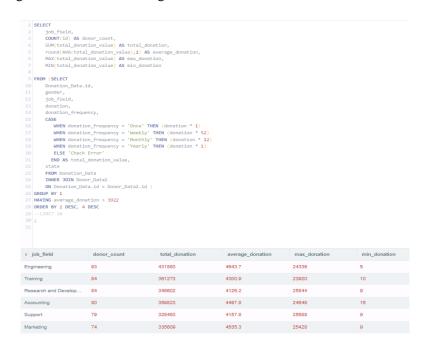


Figure 8: Top jobs by average donation

4. Number of Donors in state + job field: This is the number of donors who work in job fields and live in states where the average donation value is lower than the overall average donation value of donors.

```
SELECT

Sum(donor_count)

FROM(SELECT

State,

ROUND(AWS(total donation_value) AS total_donation,

ROUND(AWS(total donation_value),1) AS average_donation,

MAX(total_donation_value) AS max_donation

FROM(SELECT

Donation_Data.id,

gender,

job_field,

donation_frequency,

CASE

MHEN donation_frequency = 'Once' THEN (donation * 1)

MHEN donation_frequency = 'Weekly' THEN (donation * 12)

MHEN donation_frequency = 'Weekly' THEN (donation * 12)

MHEN donation_frequency = 'Yearly' THEN (donation * 12)

MHEN donation_frequency = 'Yearly' THEN (donation * 12)

ELSE 'Check Error'

END AS total_donation_value,

state

FROM Donation_Data

INNER JOIN Donor_Data2

ON Donation_Data

GROUP BY 1,2

HAVING average_donation < 3922)

I sum(donor_count)

S94
```

Figure 9: Count of Donors in low donation State and Job field

# Additional Insights

1. Top donors by state and job field: The query result is in descending order of total donation value.

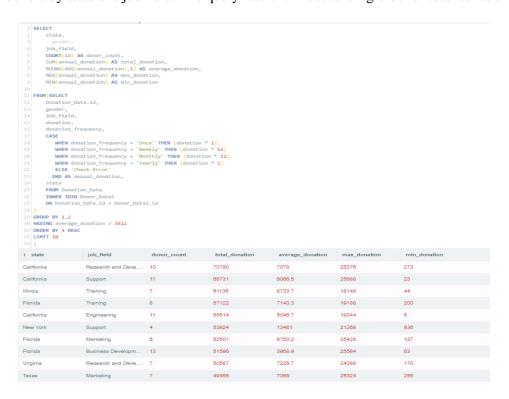


Figure 10: Top donors by State and Job field

2. Job fields by Average Donation Value: This query extracts job fields with an average donation value lower than the overall average donation value.



Figure 11: Least-donating job fields by Average Donation

3. Level of Education: This query extracts insight on the distribution of donors and donations by the donors' level of education. 'Uneducated' refers to donors with no data on university attended.

```
I SELECT

CASE WHEN university IS NOT null THEN 'Educated'

MHEN university IS null THEN 'Uneducated'

END AS Education,

COUNT(CASE WHEN university IS NOT null THEN id

END) AS donon_count,

SUM(CASE WHEN university IS NOT null THEN id

END) AS donon_count,

SUM(CASE WHEN university IS NOT null THEN total_donation_value

MHEN university IS null THEN total_donation_value

END) AS donations,

ROUND(AVG(CASE HEND university IS NOT null THEN total_donation_value

END) AS donations,

ROUND(AVG(CASE HEND university IS NOT null THEN total_donation_value

END) AS donations,

ROUND(AVG(CASE HEND university IS null THEN total_donation_value

END) AS donations,

ROUND(AVG(CASE HEND university IS null THEN total_donation_value

END) AS donations,

ROUND(AVG(CASE HEND university IS null THEN total_donation_value

END) AS donation,

donation_frequency = 'More 'THEN (donation * 1)

MHEN donation_frequency = 'Once' THEN (donation * 1)

MHEN donation_frequency = 'More 'More (donation * 12)

MHEN donation_frequency = 'More 'More (donation * 12)

MHEN donation_frequency = 'Worly' THEN (donation * 12)

MHEN donation_frequency = 'Yearly' THEN (donation * 12)

ELSE 'Check Error'

END AS total_donation_value,

university

FROM Donation_Data

INNER JOIN Donor_Data2

ON Donation_Data

INNER JOIN Donor_Data2

ON Donation_Data.id = Donor_Data2.id)

GROUP BY 1

i Education

donor_count

donations

avg_donations

Educated

748

2841228

3798.4

Uneducated

252

1081089

4290
```

Figure 12: Donor Count, Average and Total Donation Value by level of Education

## Visualizations

I gathered these insights, stored results as .csv files, then imported into Microsoft PowerBI for visualization.

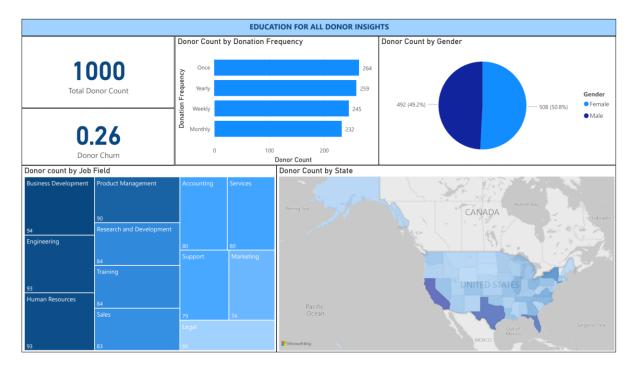


Figure 13: Donor Count Insights Visualization

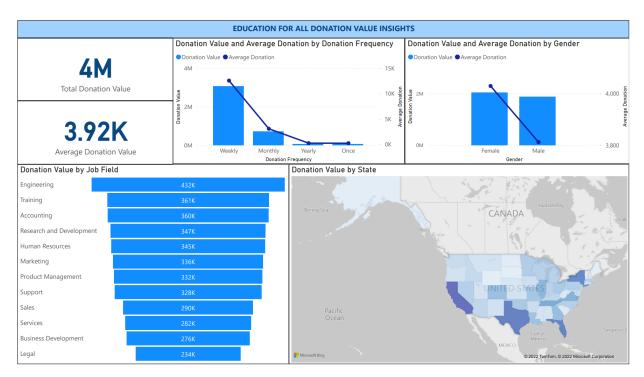


Figure 14: Donation Value Insights Visualization

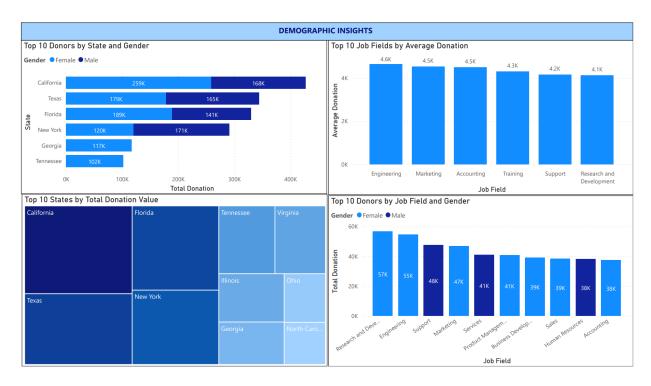


Figure 15: Donor Demographics Insight Visualization

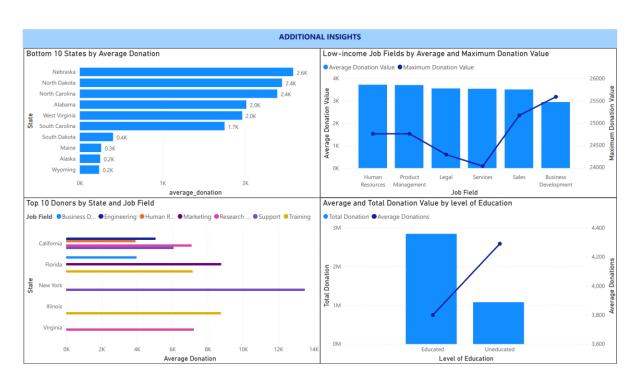


Figure 16: Additional Demographics Insight Visualization

#### **Findings & Recommendations**

#### **Findings**

- 1) The total value of donations (\$4M) to the charity in the year in review may not seem low, but it may be insufficient to fund the various projects the charity has outlined. Hence, why the charity is trying to increase donations significantly.
- 2) In the year in review, number of donors in the database is 1000, with a total donation value of \$3,922,317.
- 3) The highest donor donated a total donation value of \$25,844 this year.
- 4) Donors with a weekly frequency of donation contributed the highest percentage of the total donation value.
- 5) In order of number of donors by frequency, donors who donated just once constitute the highest percentage, followed by Donors who donate yearly (once a year). The frequency with the lowest number of donors is 'Monthly'.
- 6) 26.4% of donors donated just once, and they contributed to 1.6% of total donations.
- 7) The female gender donates more than the male gender. They have a higher donor count, donation value and average donation value than males.
- 8) It can be seen that Engineering contributes to the highest donation value of all the job fields with a value of \$431,863. Engineering also has the highest average donation value.
- High-income fields include Engineering, Training, Research and Development, Accounting, Marketing, and Support.
- 10) Top donating States include California, Texas, Florida, New York.
- 11) The states in the database with the lowest donor count are Wyoming, Maine, South Dakota, Alaska, North Dakota. By Total Donation Value, Wyoming is the lowest donating state with a donation value of \$232.
- 12) Results show that donors who work in Research & Development and live in California are the highest donors with a total donation value of \$70,780.
- 13) According to state and gender, the top donating demographic are Females and Males in California, Texas, Florida, New York, Females in Georgia, and Tennessee.

- 14) According to job field and gender, the top donating demographic are Females in Research & Development, Engineering, Marketing, Product Management, Business Development, Sales, and Accounting. Males in Support, Services and Human Resources.
- 15) Most donors donate less frequently.
- 16) We have more university-educated donors (748) than uneducated donors (252).
- 17) Donors with no university education have a higher average donation value than donors with university education.

#### Recommendations

- 1) Increase efficiency of charity's primary purpose. This boosts donor trust in the organization which may help reduce churn and increase their frequency and value of donation.
- 2) Constantly send donors personalized emails updating donors on the activities of the organization.
- 3) Investing in targeted ads, awareness programs and fundraisers to get more donors in the top states who work in high income job fields. Initiatives could also be gender, job field and location specific.
- 4) Nationwide increase in awareness programs, fundraising rallies, radio/tv advertisements. This helps to increase the total number and diversity of donors, and total donation value in the database.
- 5) Spread awareness to rural areas of top states, farm settlements, markets, etc. Areas which generally have a lower population of university educated individuals.
- 6) Prevent donor churn. Work on retention of donors getting donors to commit to the charity in some way.
- 7) Efforts should be intensified in states with the highest number of donors, to encourage them to donate higher amounts and more frequently.
- 8) Follow up with donors who donated just once via email on new projects, improvements, benefits etc. of donating to the organization.
- 9) Be open to donor complaints, compliments, recommendations, and feedback. Just like asking for donors to give, you need to be initiative-taking when requesting donors for feedback.

10) Motivate by giving away free stuff to top donors. This gift may be tailored based on the language (culture), favourite colour, or favorite movie genre.

## Conclusion

*Education For All* can achieve the objectives set by the fundraising committee if it uses insights gathered from this data to make decisions, takes actions in line with recommendations of this report.