Data Analysis Report

EDUCATION FOR ALL CHARITY ORG.

Prepared By:
Benson, Nemeke

Table of Contents

Introduction

"Education for All" is a US based charity organization whose main goal is to assist those who cannot afford education. The charity gets funding from donations made by individuals from different US states.

Project Description

This project focuses on finding useful insights from the datasets provided by the organization, to assist the fundraising team to make inform strategies to meet the following objectives for the following year;

1. To increase the number of donors.
2. To increase the donation frequency of donors.
3. To increase the value of donations to the charity.

Data Source

The dataset was gotten from the charity's database and it contains two tables 'donation data' and 'donors data'. When merged together, they contain 1000 records and 15 attributes. The dataset contains information of registered donors from 49 US states, information such as id, first name, last name, emails, gender, job field, donor's state, donations, and donation frequency. Other additional information about the donors are in the 'donors data' dataset. Before the analysis, the dataset was cleaned by checking for null records, inconsistent data types and misspellings.

Business Problem

The problem of not having enough frequent donations from individuals affects the charity's philanthropic outreach. Hence, it influences the percentage of persons in US that benefit from the charity. An efficient solution would be to apply effective fundraising strategies to meet the donations objectives for the following year.

The following are some questions asked to analyze the dataset descriptively to further understand the business problem.

1. What state has the highest and lowest amount of donors and donations?
2. What is the total number of donors and how frequent were their donations?
3. What job fields had the highest donations and donors who donated more frequently?
4. What is the ratio of female to male donors and does gender have an effect on donations?
5. How many donors acquired a tertiary education and does it inform their donations?
6. Can the data available be useful in gaining insights to help achieve the donations objectives?

Root Cause Analysis

The five whys (5whys) root cause technique was used to explore the underlying cause-and-effect of the business problem. The following why questions were asked:

1. Why is the charity not having enough donations for its philanthropic projects?

- Because the donations to the charity are not very frequent.

2. Why are the donations not very frequent?

- Because less frequent donors are more than the frequent donors are.

3. Why are less frequent donors more?

- Because there are many donors that pledged to donate but never did

4. Why did some donors pledge to donate but never donated?

- Because they probably forgot to donate.

5. Why would they not remember to donate?

- Because there probably isn't an effective reminder system to prompt the donors to donate frequently.

Insights & Visualization

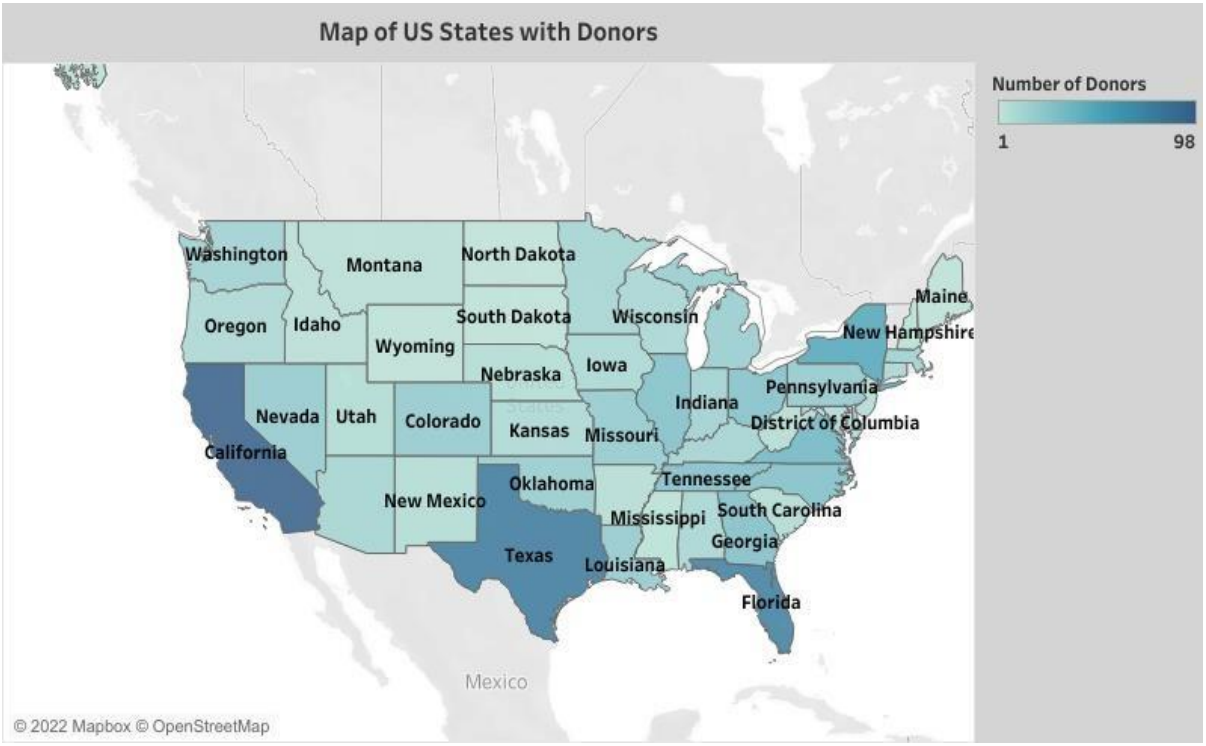
Below are insights gotten by querying the data using PostgreSQL, the query codes can be found in the appendix of this report. The insights from the queries have been graphically represented using power bi

1. States in US with Donors

   From the dataset given, it was discovered that donations to the Education for all charity came from 49 US states. That is all current US states excluding Vermont and Rhode Island.

   Figure 1 below is a map showing the different states and population density of donors in them.

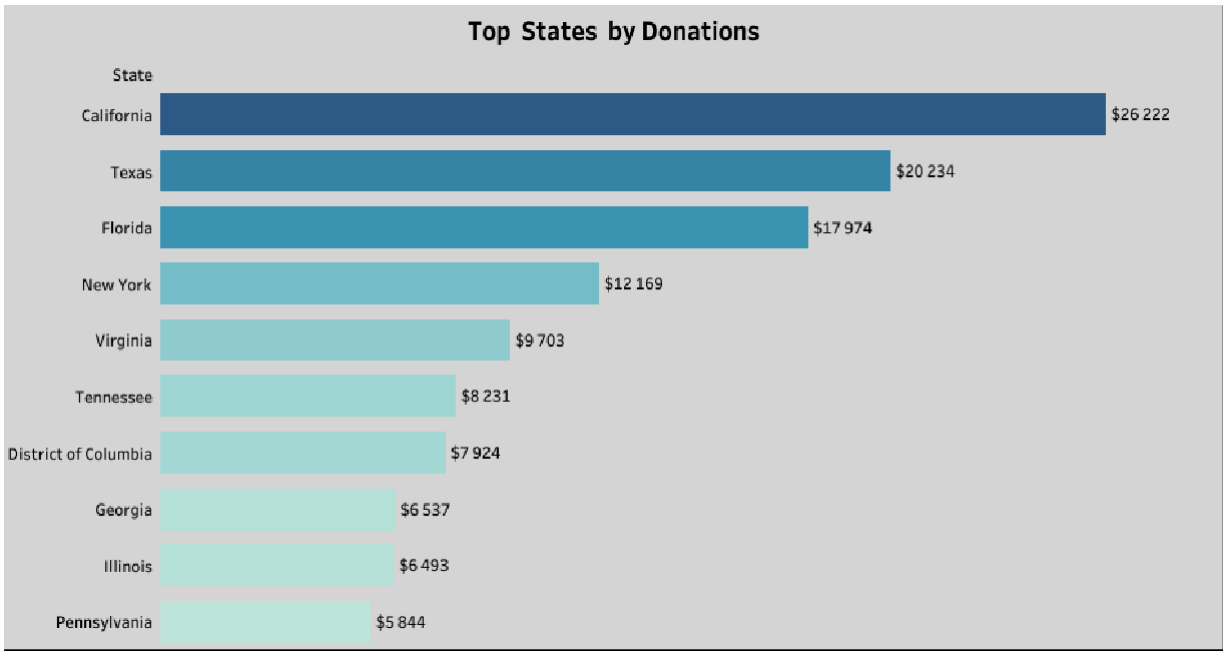
   Figure 1 Map of US States with Donors



2. Top 10 States Based on Donations

   The top states based on donations are California with $26,222 donations, Texas with $20,234, Florida with $17,974 donations followed by New York and Virginia. Figure 2 below is a graphical representation of the top 10 states and their donations.

Figure 2 Top 10 states and their Total Donations

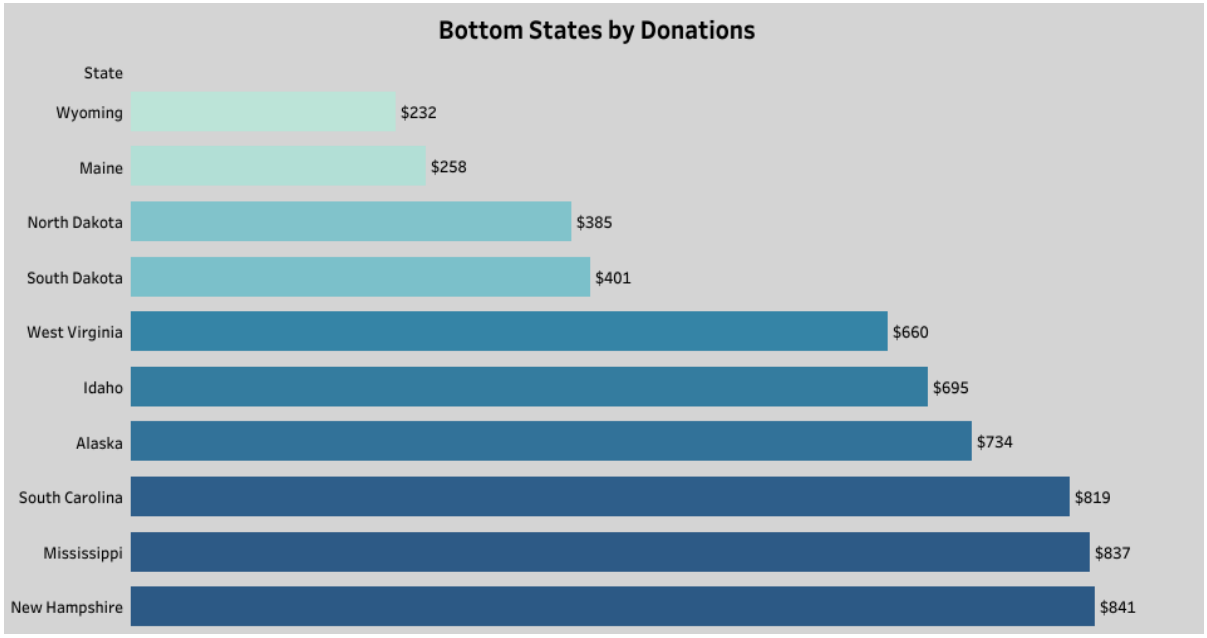


3. Bottom 10 States Based on Donations

The bottom states based on donations are Wyoming with $232, Maine with $258, North Dakota with $385 followed by South Dakota and West Virginia.

Figure 3 below is a graphical representation of the bottom 10 states and their donations.

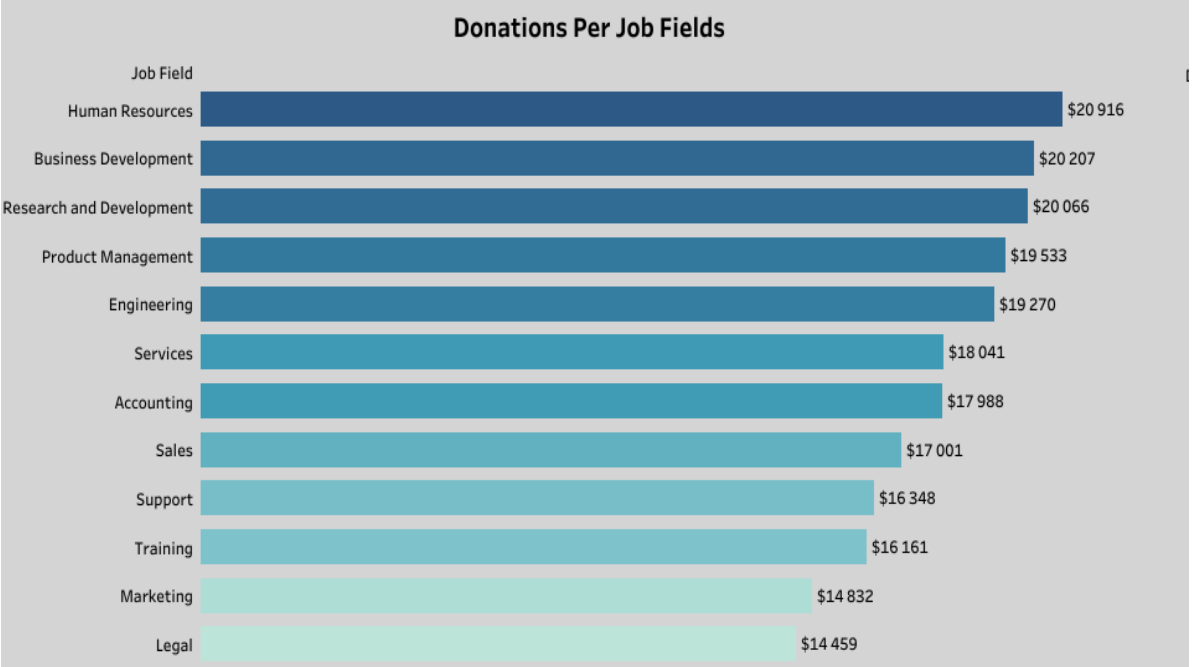
Figure 3 Bottom 10 states and their Total Donations



4. Donations Based on Job Fields

The figure below shows the total donations gotten from donors in each job field. Donations from Donors working in Human resources gave the highest subtotal donation to the charity as seen in the figure below. While donors who work in legal gave the lowest.

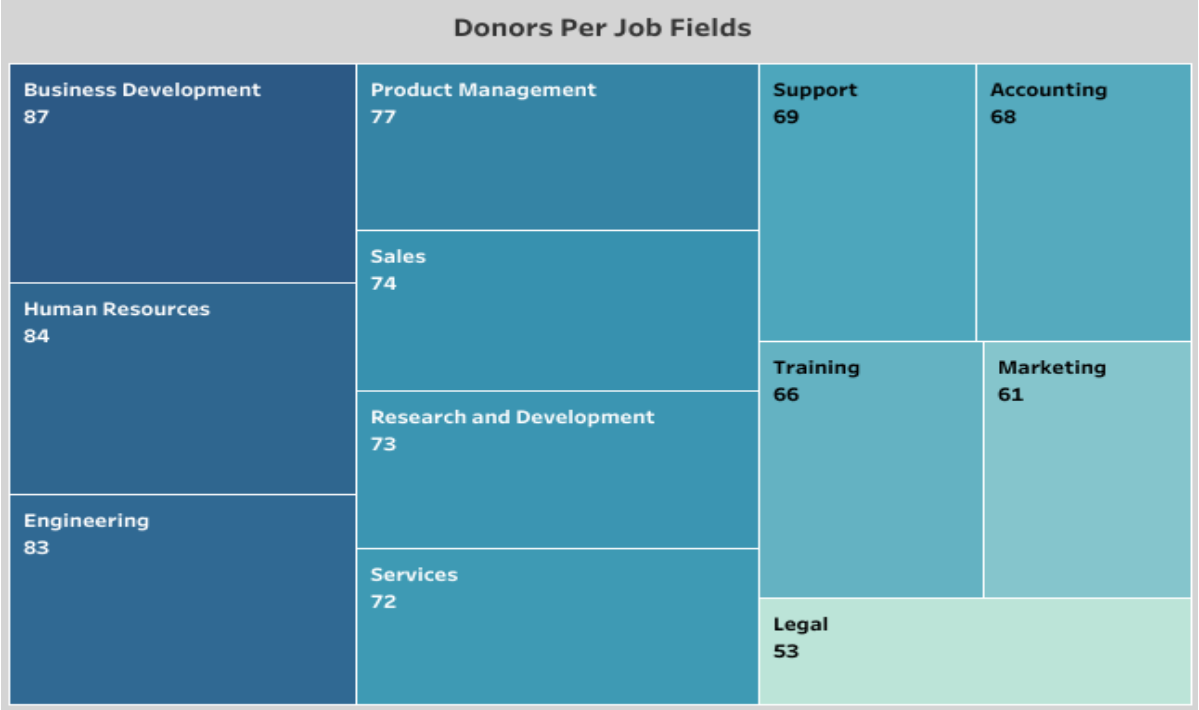
Figure 4 Job fields and their Total Donations



5. Number of Donors per Job Field

The figure below shows the total number of donors in each job field. It is observed that Business Development has the highest number of donors, followed by Human Resources and Engineering. Though Human Resources donors gave the highest donations, they are the second highest count here. It is worth observing that even though Engineering has the third highest number of donors, the field is not in the top 3 with the highest donations.

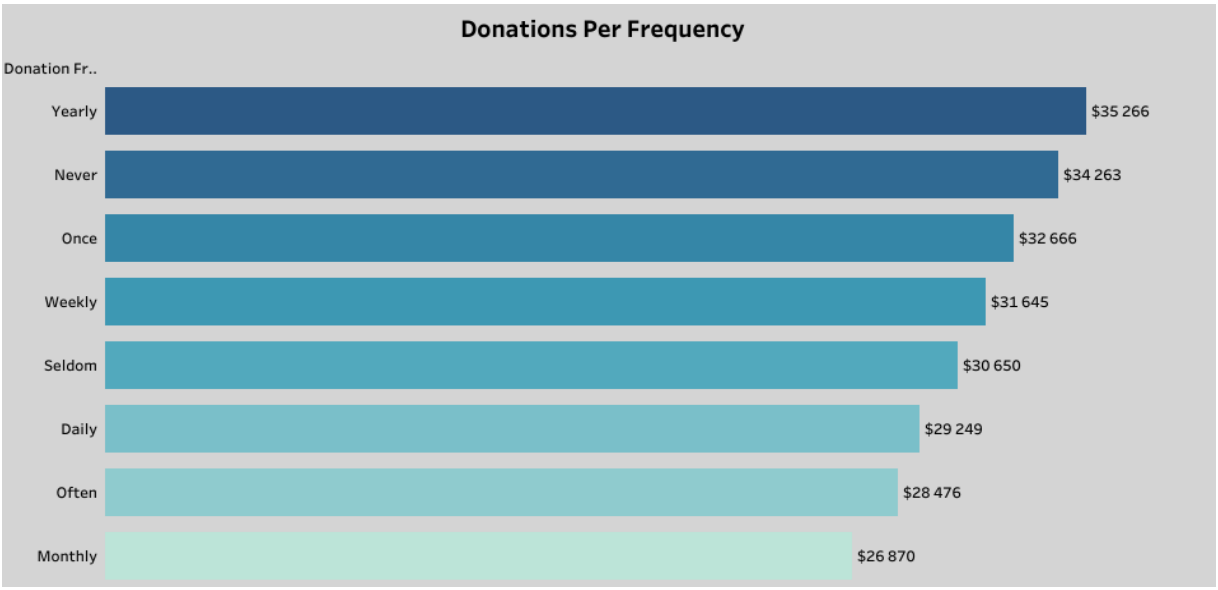
Figure 5 Number of Donors in each Job Field



6. Donations Based on Frequency

   In figure 6 below, the donations made by donors were grouped into 8 categories based on frequency. It is important to notice that the amount of donations that were never made sums up to $34,263.

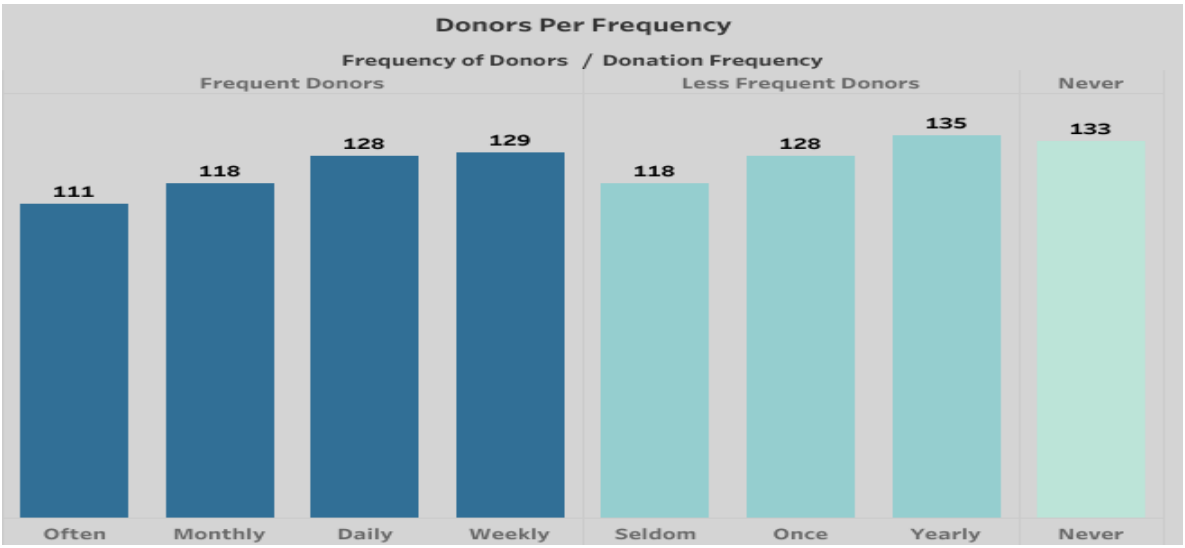
Figure 6 Donations per Frequency



7. Donors Based on Frequency

The frequency of donations was further grouped into 3 categories ‘frequent’, ‘less frequent’ and ‘Never’ as seen below. Donors who donated often, daily, weekly & monthly were grouped as frequent while donors who donated once, seldom and yearly were grouped as less frequent.

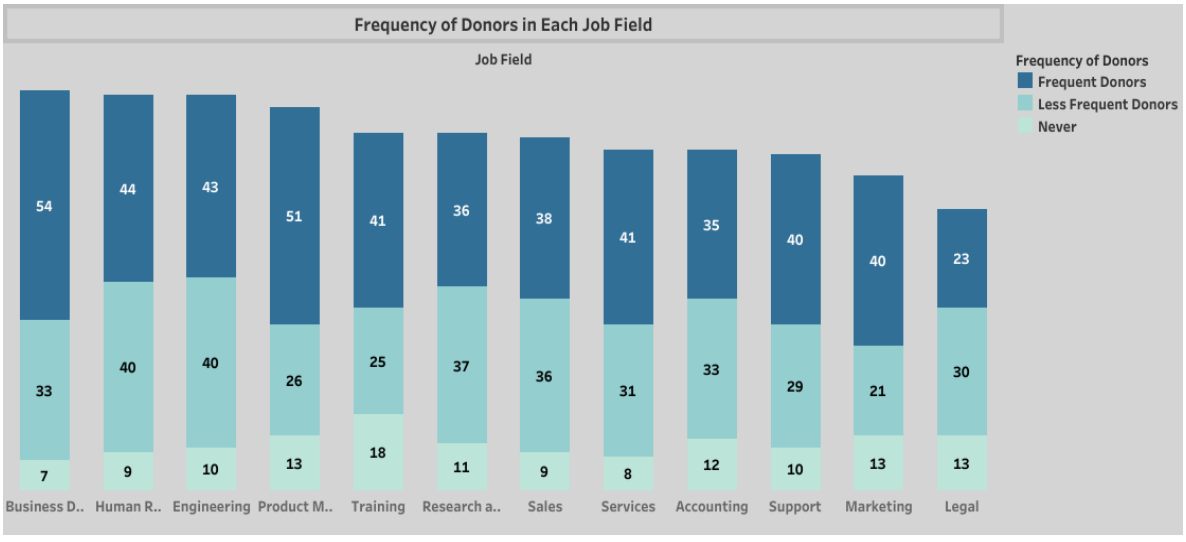
Figure 7 Donors based on Frequency



8. Frequency of Donors in Each Job Field

Figure 8 shows Business Development has the most frequent donors, Human Resources & Engineering have the highest number of less frequent donors while Training has the highest number of registered donors that never donated

Figure 8 Frequency of Donors in each Job Field



Dashboards

Figure 9 Dashboard showing information on Registered Donors

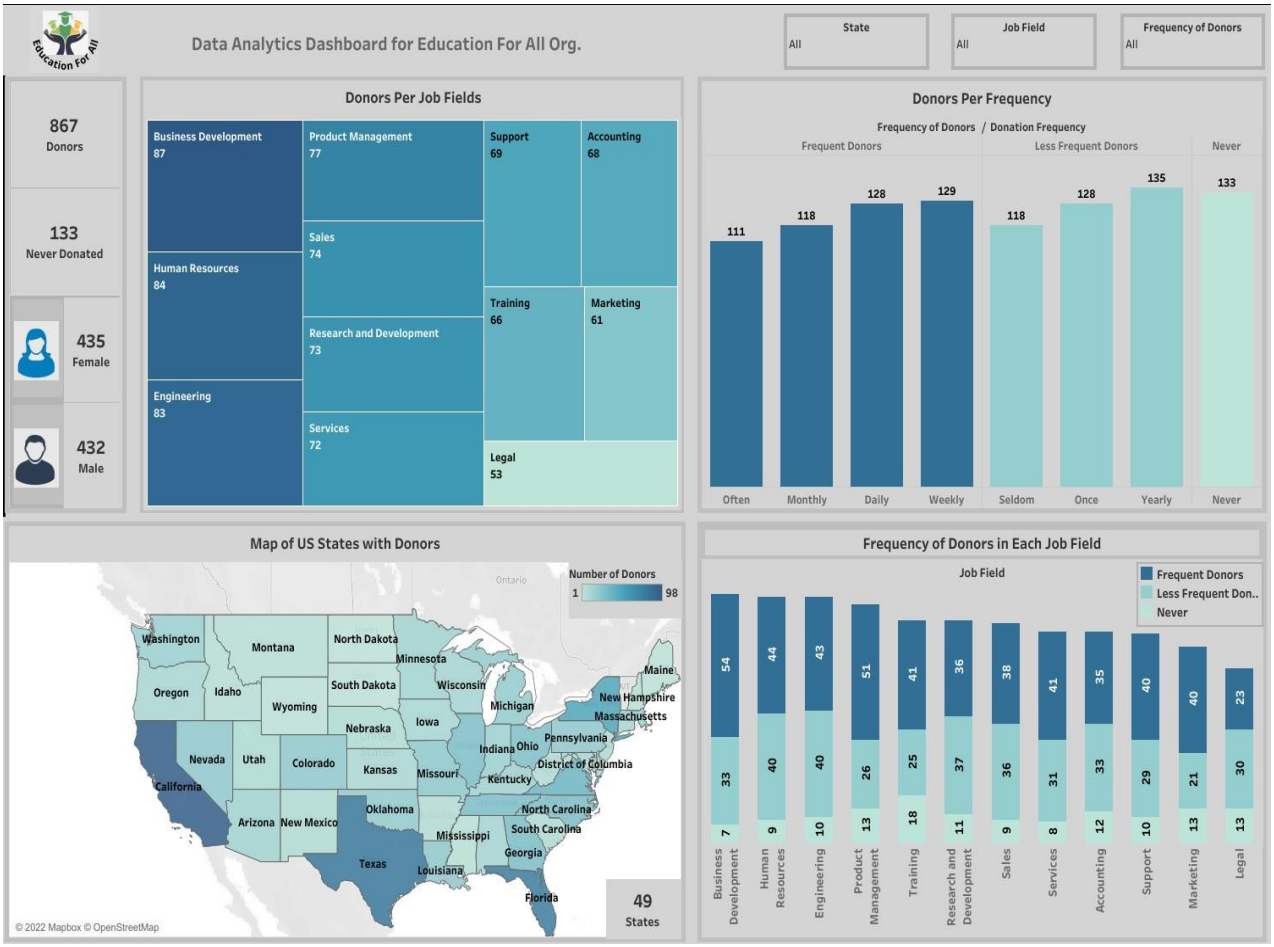
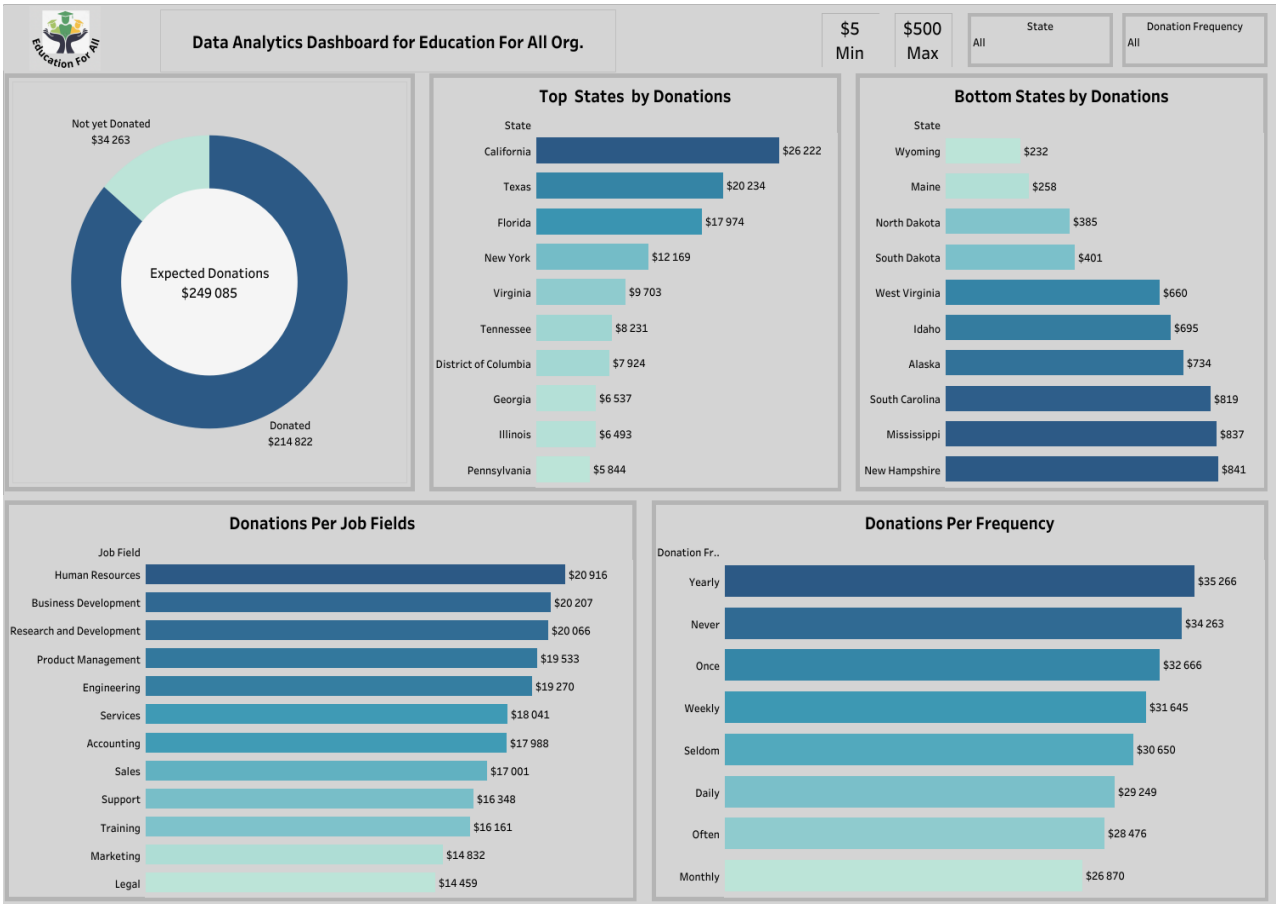


Figure 10 Dashboard showing information on Donations



Findings

After analyzing the data on the 1000 registered donors from the 49 US states, it was observed that the Charity expected $249,085 from the registered 1000 donors but only 867 donors donated and it amounted to $214,822. While 133 registered donors never donated and it amounts to a sum of $34,263, which is approximately 13.8% of the expected donations to the charity.

Donors in California gave a total of $26,222 and it was noted that California has the highest number of donors (98), followed by Texas (80). States like Wyoming, Maine, North and South Dakota have only a single donor each. Vermont &Rhode Island do not have any registered donor.

Out of 867 actual donors, 435 are females and 432 males. Though female donors are more, 48% of the donations came from females while 52% from males. Also, 468 donors acquired tertiary education and 399 did not. Only 279 registered a second language. Based on these insights it's not certain if gender, tertiary education or a second language plays a role in their donation amount.

Out of the 133 registered donors that never donated, 73 are females and they pledged to donate a total of $18,384, while 60 are males and they pledged to donate a total of $15,879. About 73 of them acquired a tertiary education but 60 did not. Only 54 of them registered a second language. It is impossible to conclude if gender, tertiary education or having a second language plays a role in them never donating, to conclude on this would require further analysis and insights on information about the donors that are currently not available in the dataset.

In an attempt to find a defining factor, the donors were categorized based on their job fields, it was observed that Business Development, Human Resources and Engineering fields have the highest number of donors to the charity, while sales has the lowest number of donors.

The donors were also categorized according to the frequency of their donations such as daily, weekly, monthly etc. The frequency of donations made by donors was further grouped into 'frequent', 'less frequent' and 'never'. The frequent category includes donors who donated daily, often, weekly and monthly, while the less frequent category includes those who donated once, seldom and yearly. The frequent donors are 486 while the less frequent donors are 381. About 18 of the donors who never donated work in training, 13 in legal, marketing and product management each. About 15 of them reside in California and Texas each. Further analysis is required to determine if job fields or state of residency plays a role in frequency and amount of donations, as there are currently not enough data provided to deduce this.

Recommendation

Based on the insights gotten from the analysis, below are some recommendations to meet the business objectives:

1. I would suggest the fundraising team could come up with strategies that would not only be applied in the current states but also in states with little to no donors to give individuals there a chance to be a part of this charity and it's cause.
2. The communications team could send out a short questionnaire via email to the registered donors to get feedback concerning the donation process. In addition, inquire from the pledgers why they haven't donated yet. In this report's appendix is a query that pulls the emails of all the registered donors that are yet to donate.
3. Communications team could set up an automated reminder email to send to donors based on the frequency intervals they registered.
4. In addition, media team could run targeted ads introducing the charity to other business developers, human resource personals, product managers, researchers and engineers since they are the top donating job fields. The charity could also consider expanding it's reach to include donors from other job fields such as education, politics, etc.
5. Further analysis and information is required to determine if job fields or state of residency plays a role in frequency and amount of donations. Also to determine if gender plays a role as female donors are more than male but their total donations is the reverse.
6. Current donation min and max value is $5 and $500 respectively, the charity could consider increasing the range of individual donation amount, and this could help increase donations.
7. I observed some important data such as information on age, annual income, family size and marital status of donors were not included in the dataset. This information on the donors would give clearer insights on external factors that may affect the amount and frequency of donations.
8. Information such as donor's favorite color, genre of movies, shirt size and car type wasn't useful in descriptively analyzing this data to understand the business problem but they may be important when coming up with targeted fundraising strategies.
9. The timeframe over which the data was collected was not provided; it wasn't clearly indicated if the data was gathered over an annual duration and the number of times each donor donated over the period was also not included. This challenge made the analysis limited. I would suggest all missing relevant information to be included in the dataset for further analysis.

Conclusion

From the insights gotten from analyzing the dataset collected, some suggestions were made to help with the business problem such as sending frequent emails to remind donors to donate to the charity as well as increasing the acceptable min and max donation amount. Also for the charity to increase its outreach to other job fields and states. These recommendations could help create more awareness for the charity "Education for all" in all the states in US and influence the frequency and amount of donations to the charity as well as number of donors therefore meeting their donations objectives for next year and thus achieving their goal to fund the education of more individuals in the US.

For further insightful analysis, the relevant missing information should be made available.

Appendix

SQL Codes

```
-- To view the content of both datasets individually
SELECT * FROM donation_data;
SELECT * FROM donor_data;
-- To view the content of both datasets merged
SELECT *
FROM donation_data dd
JOIN donor_data d
ON dd.id = d.id;
-- To view total number of registered donors and total expected donations
SELECT COUNT(id) AS donors,
		SUM(donation) AS total_donations
FROM donation_data;
-- To view total number of donors that donated and the total donations gotten by the charity
SELECT COUNT(d.id) AS donors,
		SUM(donation) AS total_donations
FROM donation_data dd
JOIN donor_data d
ON dd.id = d.id
WHERE  donation_frequency != 'Never'
--To view gender count ratio of actual donors and total donation
SELECT gender,
		COUNT(DISTINCT d.id) AS number_of_donors,
		SUM(donation) AS total_donations
FROM donation_data dd
JOIN donor_data d
```

```
ON dd.id = d.id
WHERE  donation_frequency != 'Never'
GROUP BY 1
ORDER BY 2 DESC;
```

--To view the gender count ratio of donors that pledged to donate but never did.

```
SELECT gender,
                COUNT(DISTINCT d.id) AS number_of_donors,
                SUM(donation) AS total_donations
FROM donation_data dd
JOIN donor_data d
ON dd.id = d.id
WHERE  donation_frequency = 'Never'
GROUP BY 1
ORDER BY 2 DESC;
```

-- To view all the US states with donors

```
SELECT DISTINCT state
FROM donation_data;
```

-- To view number of donors and total donation per state in descending order

```
SELECT state,
                COUNT(DISTINCT d.id) AS number_of_donors,
                SUM(donation) AS total_donations
FROM donation_data dd
JOIN donor_data d
ON dd.id = d.id
WHERE  donation_frequency != 'Never'
GROUP BY 1
ORDER BY 3 DESC;
```

-- To view number of donors and total donation per job field in descending order

```
SELECT job_field,
                COUNT(DISTINCT d.id) AS number_of_donors,
                SUM(donation) AS total_donations
```

```
FROM donation_data dd
JOIN donor_data d
ON dd.id = d.id
WHERE  donation_frequency != 'Never'
GROUP BY 1
ORDER BY 3 DESC;
```

-- To view number of donors and total donation per donation frequency in descending order

```
SELECT donation_frequency,
                COUNT(DISTINCT id) AS number_of_donors,
                SUM(donation) AS total_donations
FROM donation_data dd
JOIN donor_data d
ON dd.id = d.id
GROUP BY 1
ORDER BY 2 Desc;
```

-- Using subqueries to view donors and their total donations after further grouping the frequency of donations into 2 categories 'frequent' and 'less frequent'

```
SELECT T1.frequency,
                COUNT(T2.id) AS number_of_donors,
                SUM(T2.donation) AS total_donations
FROM (SELECT job_field,
                d.id as id,
                donation,
                donation_frequency,
        CASE
        WHEN donation_frequency = 'Never'
        OR donation_frequency ='Once'
        OR donation_frequency = 'Seldom'
        OR donation_frequency = 'Yearly' THEN 'less frequent'
        ELSE 'frequent'
        END AS frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id) T1
```

```
JOIN
(SELECT job_field,
                d.id as id,
                donation,
                donation_frequency,
        CASE
        WHEN donation_frequency = 'Never'
        OR donation_frequency ='Once'
        OR donation_frequency = 'Seldom'
        OR donation_frequency = 'Yearly' THEN 'less frequent'
        ELSE 'frequent'
        END AS frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id) T2
ON T1.id = T2.id
GROUP BY 1
ORDER BY 3 DESC;
```

-- To view the gender ratio of frequent and less frequent donors and total donations

```
SELECT T1.gender,
                T1.frequency,
                COUNT(T2.id) AS number_of_donors,
                SUM(T2.donation) AS total_donations
FROM (SELECT gender,
                d.id as id,
                donation,
                donation_frequency,
        CASE
        WHEN donation_frequency = 'Never'
        OR donation_frequency ='Once'
        OR donation_frequency = 'Seldom'
        OR donation_frequency = 'Yearly' THEN 'less frequent'
```

```
        ELSE 'frequent'
        END AS frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id) T1
JOIN
(SELECT gender,
                d.id as id,
                donation,
                donation_frequency,
        CASE
        WHEN donation_frequency = 'Never'
        OR donation_frequency ='Once'
        OR donation_frequency = 'Seldom'
        OR donation_frequency = 'Yearly' THEN 'less frequent'
        ELSE 'frequent'
        END AS frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id) T2
ON T1.id = T2.id
GROUP BY 1, 2
ORDER BY 1;
```

-- To view the number of frequent and less frequent donors and total donations from the different job fields

```
SELECT T1.job_field, T1.frequency,
                COUNT(T2.id) AS number_of_donors,
                SUM(T2.donation) AS total_donations
FROM (SELECT job_field,
                d.id as id,
                donation,
                donation_frequency,
        CASE
```

```
        WHEN donation_frequency = 'Never'
        OR donation_frequency ='Once'
        OR donation_frequency = 'Seldom'
        OR donation_frequency = 'Yearly' THEN 'less frequent'
        ELSE 'frequent'
        END AS frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id) T1
JOIN
(SELECT job_field,
                d.id as id,
                donation,
                donation_frequency,
        CASE
        WHEN donation_frequency = 'Never'
        OR donation_frequency ='Once'
        OR donation_frequency = 'Seldom'
        OR donation_frequency = 'Yearly' THEN 'less frequent'
        ELSE 'frequent'
        END AS frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id) T2
ON T1.id = T2.id
GROUP BY 1, 2
ORDER BY 1;
```

-- To get information on the top 10 frequent donors based on the value of their donation

```
SELECT T2.id AS id,
                T1.name,
                T1.gender,
                T1.state,
```

```
                T1.university,
                T1.job_field,
                T1.frequency,
                T2.donation AS donation
FROM (SELECT job_field,
                concat (first_name,' ',last_name) AS name,
                d.id AS id,
                gender,
                state,
                university,
                donation,
                donation_frequency,
        CASE
        WHEN donation_frequency = 'Never'
        OR donation_frequency ='Once'
        OR donation_frequency = 'Seldom'
        OR donation_frequency = 'Yearly' THEN 'less frequent'
        ELSE 'frequent'
        END AS frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id) T1
JOIN
(SELECT job_field,
                concat (first_name,' ',last_name) AS name,
                d.id AS id,
                donation,
                donation_frequency,
        CASE
        WHEN donation_frequency = 'Never'
        OR donation_frequency ='Once'
        OR donation_frequency = 'Seldom'
        OR donation_frequency = 'Yearly' THEN 'less frequent'
```

```
        ELSE 'frequent'
        END AS frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id)T2
ON T1.id = T2.id
WHERE T1.frequency = 'frequent'
ORDER BY 8 DESC
LIMIT 10;
```

--To view information on donors that pledged to donate but never did.

```
SELECT T2.id AS id,
                T1.name,
                T1.gender,
                T1.email,
                T2.state,
                T2.university,
                T2.job_field,
                T1.donation
FROM(SELECT d.id as id,
                concat (first_name,' ',last_name) AS name,
                gender,
                email,
                donation,
                donation_frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id)T1
JOIN (SELECT d.id as id,
                state,
                university,
                job_field
        FROM donation_data dd
        JOIN donor_data d
```

```
        ON dd.id = d.id)T2
ON T1.id = T2.id
WHERE  T1.donation_frequency = 'Never'
ORDER BY 2;
```

--To view the gender count ratio of donors that pledged to donate but never did but acquired a tertiary education.

```
SELECT T1.gender, COUNT(T1.university) AS number_of_donors,
                SUM(T1.donation) AS total_donations
FROM (SELECT job_field,gender, university,
                d.id as id,
                donation,
                donation_frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id
        WHERE university is NOT NULL) T1
JOIN (SELECT job_field,gender, university,
                d.id as id,
                donation,
                donation_frequency
        FROM donation_data dd
        JOIN donor_data d
        ON dd.id = d.id
        WHERE university is NOT NULL) T2
ON T1.id = T2.id
WHERE T1.donation_frequency = 'Never'
GROUP BY 1;
```