

Data Analysis portfolio

I'm an aspiring data analyst transiting my career from an auditor with great level of creative and analytical skills, problem solving and critical thinking. I have the ability to pay attention to detail, adapt and learn things faster. I'm open in working in a team or as an individual.

Professional Background

I've worked as an auditor for Office of State Auditor General, Lagos state (OSAG). I've also worked as an IT auditor in Economic Community of West African State (ECOWAS). During my time as an auditor, I've engaged in filing, recording, compiling, and transmitting financial records. I have great experience in detecting and preventing fraudulent activities. I've also developed and prepared detailed reports of audits findings. Therefore I'm a person that pay great attentions to detail, analyze and suggest solutions to problems.

At this stage of my life, I'm without a job but I'm using this time to train as a data analyst at DataCamp to ensure I've enough knowledge and practices to smoothly transit my career into a full time data analyst.

I believe that having graduated as an Accounting student, having some experience in collecting data, analyzing, calculating and preparing detailed report as an auditor as well as taking steps to further educate myself in data analysis, those will be helpful to me to work as a data analyst.

I'm for open for a new challenges and discoveries and will be thrilled to contribute my skills. Having said that, I'm a hard worker and always willingly learn, improve and explore in a bid to ensure that I contribute to achieving the organization goals and objective.

"There are no secrets to success. It is the result of preparation, hard work, and learning from failure." - Colin Powell

TABLE OF CONTENT

Professional Background	1
Introduction.....	3
Root Cause Analysis Process.....	3
Insights from the Analysis.....	5
Data analysis and visualization using tableau	12
Findings and Recommendation.....	15
Conclusion.....	21

Introduction

As a Data Analyst working for the charity called EDUCATION FOR ALL. I was asked by the Head of Fundraising to present the data on donor insights and donation rates.

Within the Fundraising team, my objectives are to:

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

My team is having a fundraising strategy meeting for the following year and I must present insights from the donation and donor data given to me to inform them and increase donations.

I used the data sets EFO_Donation_Data and EFO_Donor_Data given by the organization to answer the business problem.

I applied SQL commands to analyze data: JOIN, ORDER BY, WHERE, BETWEEN, AND, OR, SUM(), COUNT(), AVG(), GROUP BY, HAVING.

Also, I used Root Cause Analysis to understand the problem and ask right questions.

As a result, I have found out crucial insights of provided data sets, prepared visualizations, and report for my team.

Root Cause Analysis Process

The business problem is the inability to self-generate income and the inconsistency of donors in maintaining donation to sustain the goals of the charity organization.

In order to get to the root of the problem, they must rake in more donors, try to convince them to insert a standing order for their donation so they can donate more regularly.

I must present my report in two weeks.

To understand the problem, I had to analyze the data bases of Donors and Donations given to me. Also, I should present vital numbers and visualizations of the data sets.

I asked some questions to dig the problem deeper:

- I. How many donors we have in existing database including the male donors and female donors?
- II. What is the amount of their donations including the male donors and female donors?
- III. What is the average of their donations including the male donors and female donors?
- IV. What is the frequency of the donations?
- V. Who are the top 20 donors?
- VI. Does amount of donations depend on gender, car, university degree, job field?

Also, to understand this fully, I realized that I should analyze and figure out what is happening, why some people donate regularly and some of them do not, whether there's a correlation between having university education and donation and a correlation between job fields and donations. I had to specify the main symptoms and trends of the dataset.

Additionally, I applied Root Cause Analysis to ask:

1. Why some people are not interested in donating?
2. Why we do not advise our donors to add a standing order to their bank accounts so as to donate regularly?
3. Why we do not ask our donors whether they feel good or feel a sense of happiness when donating?
4. Why we do not add the economic trend to our dataset to determine if donor's income was affected and how they vary to donations?
5. Why male donors donated more than female donor despites female being the most donors in our dataset?

Insights from the Analysis

I was provided with 2 relational databases such as: EFO_Donation_Data and EFO_Donor_Data to answer the business problem. SQLite Database Management System was used to find out main insights.

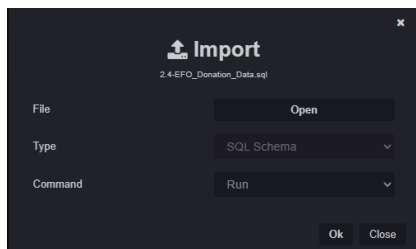
Donation Dataset includes such data:

- Id
- First name
- Last Name
- Email
- Gender
- Job field
- Donation
- State
- Shirt Size

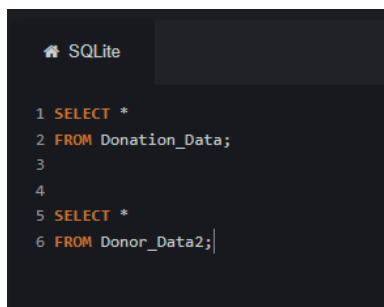
Donor Dataset includes such data:

- Id
- Donation frequency
- University
- Car
- Second language
- Favourite colour
- Movie genre

Both data sets were imported into SQLite online.

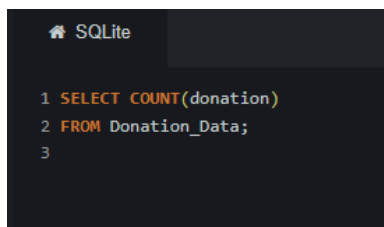


I used the **SELECT** statement to fetch the data from a database.



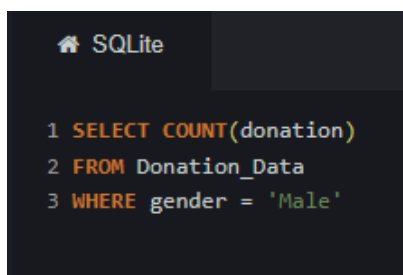
```
SQLite
1 SELECT *
2 FROM Donation_Data;
3
4
5 SELECT *
6 FROM Donor_Data2;
```

I used **COUNT()** function to get the total number of donors. It can be used like this:



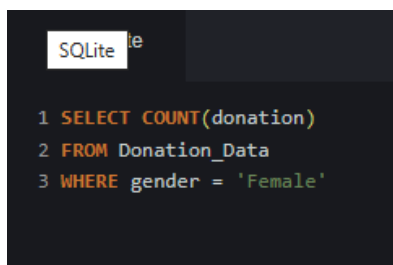
```
SQLite
1 SELECT COUNT(donation)
2 FROM Donation_Data;
3
```

I used the **COUNT()** function to get the total number of male that are donors. **WHERE** function was specifically used to filter the Male.



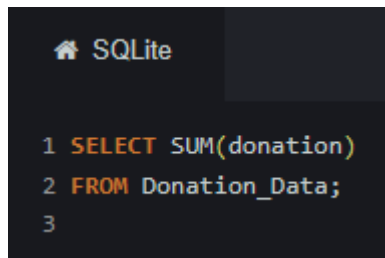
```
SQLite
1 SELECT COUNT(donation)
2 FROM Donation_Data
3 WHERE gender = 'Male'
```

I used the **COUNT()** function to get the total number of male that are donors. **WHERE** function was specifically used to filter the Female.



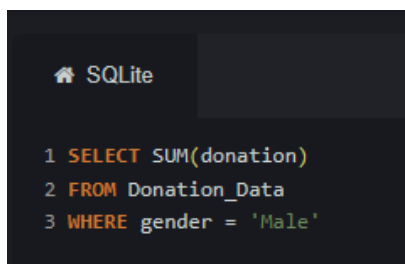
```
SQLite
1 SELECT COUNT(donation)
2 FROM Donation_Data
3 WHERE gender = 'Female'
```

To know the total sum of the donation, the function used is **SUM()** function and it can be used like this:



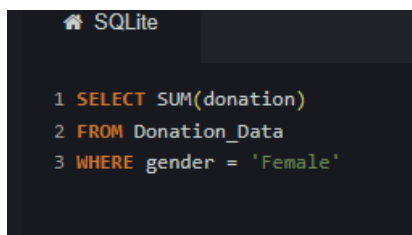
```
SQLite
1 SELECT SUM(donation)
2 FROM Donation_Data;
3
```

I used the **SUM()** function to get the total sum of males that donated. **WHERE** function was specifically used to filter the Male:



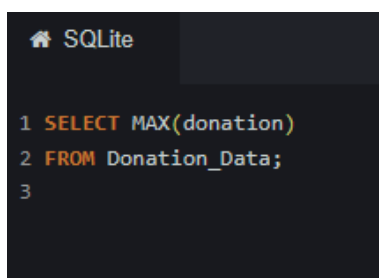
```
SQLite
1 SELECT SUM(donation)
2 FROM Donation_Data
3 WHERE gender = 'Male'
```

I used the **SUM()** function to get the total sum of males that donated. **WHERE** function was specifically used to filter the Female:



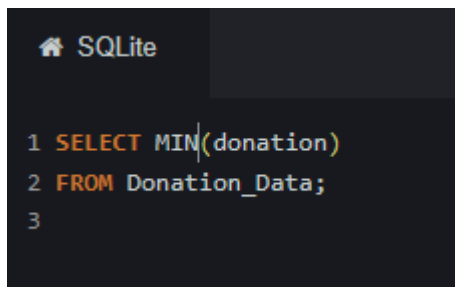
```
SQLite
1 SELECT SUM(donation)
2 FROM Donation_Data
3 WHERE gender = 'Female'
```

The maximum or the largest amount of donation can be gotten with **MAX()** function like this:



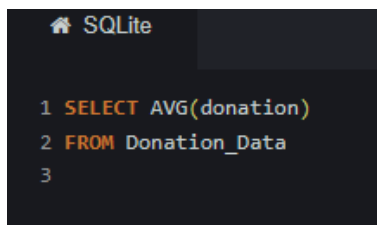
```
SQLite
1 SELECT MAX(donation)
2 FROM Donation_Data;
3
```

The minimum or the smallest amount of donation can be gotten with **MIN()** function like this:



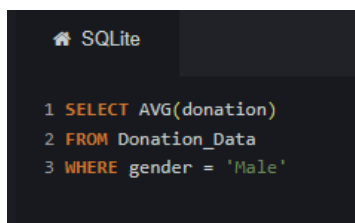
```
SQLite
1 SELECT MIN(donation)
2 FROM Donation_Data;
3
```

The **AVG()** function was used to calculate the average amount of donation and it's used as displayed below:



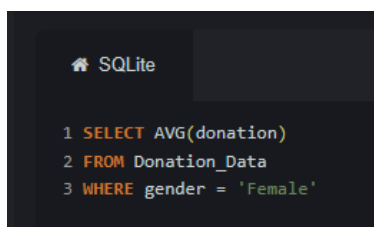
```
SQLite
1 SELECT AVG(donation)
2 FROM Donation_Data
3
```

To get the average amount of donation of Male, the **AVG()** function with **WHERE** function was used to carry out the operation and it used like this:



```
SQLite
1 SELECT AVG(donation)
2 FROM Donation_Data
3 WHERE gender = 'Male'
```

To get the average amount of donation of Female, the **AVG()** function with **WHERE** function was used to carry out the operation and it used like this:



```
SQLite
1 SELECT AVG(donation)
2 FROM Donation_Data
3 WHERE gender = 'Female'
```

Here, we want to know the top 5 job_field having over 50 donors. The **HAVING()** function is used this way. The **ORDER BY** with **DESC** was used to arrange the donor are working starting from the highest to the lowest.


```
SQLite

1 SELECT job_field, COUNT(*)
2 FROM Donation_Data
3 GROUP BY job_field
4 HAVING COUNT (*) >50
5 ORDER BY COUNT(*) DESC
6 LIMIT 5;
```

Top 20 donors with positive ratio

INNER JOIN function can be used to join different columns from different table together. I also used WHERE function with LIMIT of 20 to determine the Top 20 donors monthly.

```
SQLite

1 SELECT Donation_Data.donation, Donation_Data.gender, Donation_Data.state,Donor_Data2.donat
2 FROM Donation_Data
3 LEFT JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 WHERE donation_frequency = 'Monthly'
6 AND donation BETWEEN 350 AND 500
7 ORDER BY donation DESC
8 LIMIT 20;
```

Top 20 donors with bad ratio

In this case, it's similar to the above but instead of the above **WHERE** donation **BETWEEN** 350 **AND** 500, It changed to **BETWEEN** 0 **AND** 50 here.

```
SQLite

1 SELECT Donation_Data.gender, Donation_Data.donation, Donor_Data2.donation_frequency, car, sh
2 FROM Donation_Data
3 INNER JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 WHERE donation BETWEEN 0 AND 50
6 AND donation_frequency = 'Monthly'
7 ORDER BY donation DESC
8 LIMIT 20;
```

Donation Frequency

I also used **INNER JOIN** function to know the amount of donation done once, weekly, monthly and yearly.

```
1 SELECT Donation_Data.donation, Donor_Data2.donation_frequency, SUM(donation)
2 FROM Donation_Data
3 INNER JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 AND donation_frequency = 'Once';
6
```

```
1 SELECT Donation_Data.donation, Donor_Data2.donation_frequency, SUM(donation)
2 FROM Donation_Data
3 INNER JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 AND donation_frequency = 'Weekly';
6
```

```
1 SELECT Donation_Data.donation, Donor_Data2.donation_frequency, SUM(donation)
2 FROM Donation_Data
3 INNER JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 AND donation_frequency = 'Monthly';
6
```

```
1 SELECT Donation_Data.donation, Donor_Data2.donation_frequency, SUM(donation)
2 FROM Donation_Data
3 INNER JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 AND donation_frequency = 'Yearly';
6
```

I used the **INNER JOIN** or **JOIN** Function to check many donors have cars and can be gotten like this:

```
1 SELECT COUNT(*)
2 FROM Donation_Data
3 INNER JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 WHERE car IS NOT NULL;
```

I used the command function **LEFT JOIN** on Donation_Data and Donor_Data2 by the combination of **ON** to analyze and understand top 20 donors with university education donated \$350 or more using the **WHERE** function, **AND** function, **IS NOT NULL** function, **ORDER BY** function and **LIMIT** function. It written like this:

With university education:

```
1 SELECT Donation_Data.donation, Donation_Data.gender, Donation_Data.state, Donation_Data.job_
2 FROM Donation_Data
3 JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 WHERE donation >= 350
6 AND university IS NOT NULL
7 ORDER BY donation DESC
8 LIMIT 20;
```

Here, I who are the used the command function **LEFT JOIN** function on Donation_Data and Donor_Data2 by the **ON** function to analyze and understand top 20 donors without university education donated \$350 or more using the **WHERE** function, **AND** function, **IS NULL** function, **ORDER BY** function and **LIMIT** function. It written like this:

Without university education:

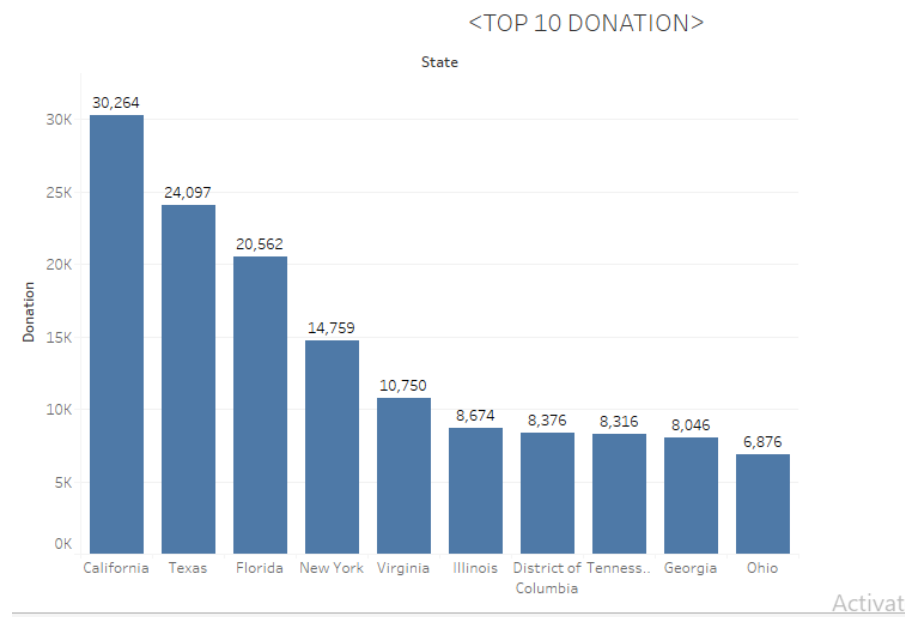
```
SQLite
1 SELECT Donation_Data.donation, Donation_Data.gender, Donation_Data.state, Donation_Data.job_
2 FROM Donation_Data
3 JOIN Donor_Data2
4 ON Donation_Data.id = Donor_Data2.id
5 WHERE donation >= 350
6 AND university IS NULL
7 ORDER BY donation DESC
8 LIMIT 20;
```

DATA ANALYSIS AND VISUALIZATION USING TABLEAU

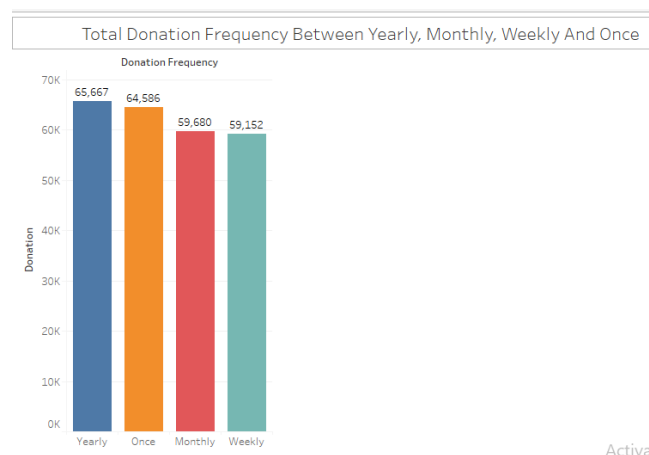
Tableau is a very powerful tool used for analyzing and visualizing data. It help in translating queries and data into visual forms, managing mega size data, help in designing and publishing interactive dashboards and visualizations.

Below are the visualizations that can be used in conveying messages.

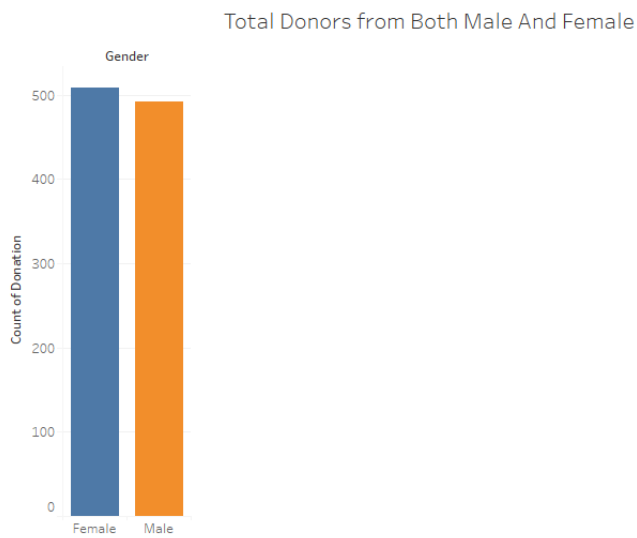
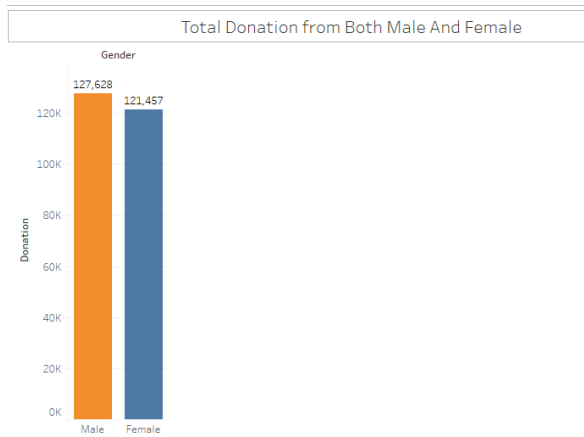
This visualization below show the top 10 donation from each state which California donated the most followed by Texas and Florida.



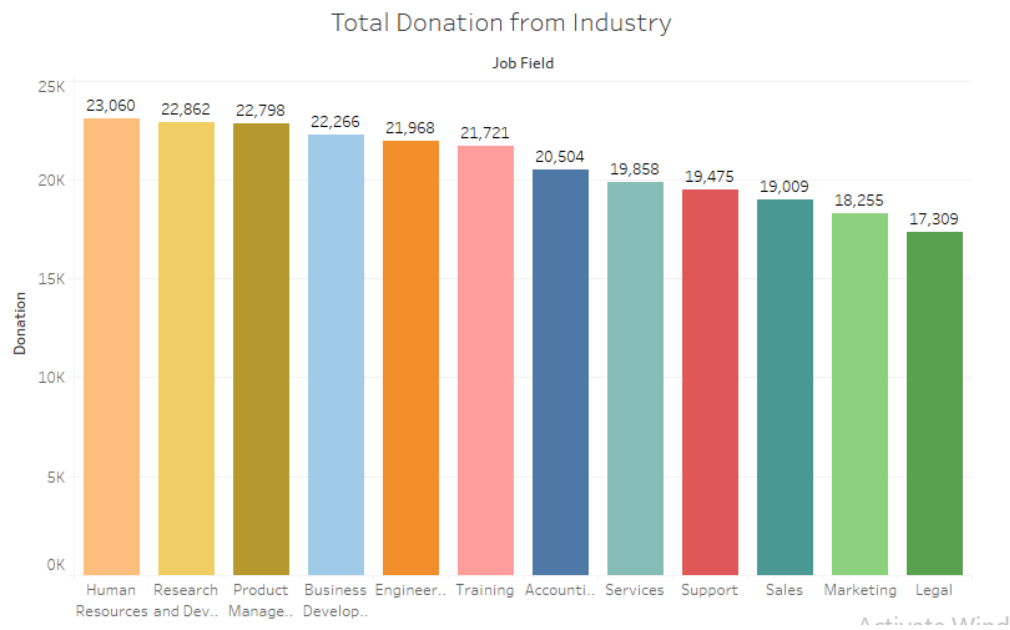
This visualization clearly show that donors prefer to donate yearly and once, with the least donors donating weekly.



Female are the most donors between the gender but despite male being the least donors they donated more amount of donation than the female as shown in the visualization below.



The next visualization show that donors from Human resources, Research and Development and Product Management donated the most with Human Resources being the highest.



Findings and Recommendations

After analyzing the data set given, I found out that;

- The total number of donors we have in the database is 1000
- The sum of donations we collected is \$249085
- The average sum of donations is \$ 249.09
- The largest amount of donations is \$500
- The smallest amount of donation is \$5
- The sum of donations from male is \$127,628
- The sum of donation from female is \$121,457
- The number of donors with car is 939 while the number of donors without car is 61
- The number of donors with university education is 748 while the number of donors without university education is 252
- The number of donors that are male is 492 while the number of donors that are female is 508

Frequency and amount of Donation

Donation	donation_frequency	Sum(donation)
292	Once	64586
28	Weekly	59512
178	Monthly	59680
255	Yearly	65667

We were given a data base of 1000 donors of which only 753 out of 1000 are active donors. 247 donors are not active which can be gotten from subtracting the active donors from 1000 donors. It means that 247 donors didn't donate at all. Also, we can see that donors prefer to donate yearly or once. Therefore, we should put more effort in ensuring we reach out to the inactive donors to find how why they are not donating and also ensure we increase the frequency of the donors by donating as least monthly.

Top 5 job field donors

job_field	COUNT(*)
Business Development	94
Human Resources	93
Engineering	93
Product Management	90
Training	84

Top 20 donors who donates monthly between \$350 and \$500

donation	Gender	State	donation_frequency	job_field	Car	university
500	Male	New York	Monthly	Product Management	Pontiac	Leithgoe
498	Male	Wisconsin	Monthly	Sales	Volvo	Trotton
494	Female	California	Monthly	Product Management	Maybach	Coates
494	Male	California	Monthly	Human Resources	Dodge	Baumber
493	Male	Maryland	Monthly	Product Management	Infiniti	Armatidge
491	Female	New Mexico	Monthly	Engineering	Mercedes-Benz	Blackborn
490	Female	Kentucky	Monthly	Engineering	Mercedes-Benz	Heinert
489	Male	California	Monthly	Training	Audi	Menci
483	Female	Connecticut	Monthly	Product Management	BMW	Gorch
483	Female	Pennsylvania	Monthly	Legal	Mazda	Null
482	Male	California	Monthly	Sales	Lamborghini	Baford
480	Male	Tennessee	Monthly	Human Resources	Mercedes-Benz	Null
480	Female	Arkansas	Monthly	Engineering	Nissan	Null

478	Male	Massachusetts	Monthly	Services	Ford	Pardoe
474	Male	Utah	Monthly	Engineering	Audi	Finlayson
466	Male	Virginia	Monthly	Training	Toyota	Hunting
465	Male	California	Monthly	Research and Development	Kia	Morhall
462	Female	Pennsylvania	Monthly	Support	Mercedes-Benz	Null
460	Male	Texas	Monthly	Product Management	Ford	Bushill
458	Male	California	Monthly	Support	Subaru	Mallall

Top 20 negative donors who donates monthly between \$0 and \$50

Donation	Gender	State	donation_frequency	job_field	Car	University
50	Male	New York	Monthly	Accounting	Mazda	Null
46	Female	Arizona	Monthly	Services	Volkswagen	Dundon
45	Male	Nebraska	Monthly	Human Resources	Oldsmobile	Frankis
44	Female	Georgia	Monthly	Product Management	Land Rover	Bramble
42	Male	Florida	Monthly	Engineering	Hyundai	Null
41	Male	North Carolina	Monthly	Marketing	Chevrolet	Millgate
40	Male	District of Columbia	Monthly	Legal	Cadillac	Dandison
39	Male	Alabama	Monthly	Services	Toyota	Caley
37	Female	Texas	Monthly	Product Management	Volvo	Herries
37	Male	Florida	Monthly	Accounting	Dodge	Murison
35	Female	Florida	Monthly	Support	Ford	Rude
34	Female	Pennsylvania	Monthly	Accounting	Saab	Capnor
33	Female	Florida	Monthly	Services	Ford	Null

30	Female	New York	Monthly	Research and Development	Chevrolet	Null
29	Female	New Hampshire	Monthly	Services	Honda	Null
27	Female	Colorado	Monthly	Research and Development	Toyota	Giannasi
25	Female	Washington	Monthly	Marketing	BMW	Null
25	Male	North Carolina	Monthly	Accounting	Isuzu	Cayley
22	Male	Tennessee	Monthly	Sales	Dodge	Null
19	Male	Missouri	Monthly	Human Resources	Cadillac	Null

The analysis shown that donors from different states vary and their donation depends more on their line of job. For instance, most donors donated higher amount of donations from California than other states but donors from Product management donated the most monthly. We can see that donors working in product management are more likely to donate higher than any other job field. Also, those having expensive car or not doesn't have any relationship with the amount they donate. Lastly, most that donated have university education whether they donate high or low. Note that this donation frequency is analyzed monthly.

Top 20 donors with University Education

Number	Donation	gender	State	job_field	University	car
1	500	Male	Michigan	Support	Walasik	Honda
2	500	Male	New York	Product Management	Leithgoe	Pontiac
3	499	Female	Virginia	Legal	Sparhawk	Ford
4	499	Female	Delaware	Sales	Antoszewski	Ford
5	498	Male	Wisconsin	Sales	Trotton	Volvo
6	497	Male	New York	Research and Development	Rockcliffe	Maserati
7	494	Male	California	Support	Cominetti	Kia
8	494	Female	California	Product Management	Coates	Maybach
9	494	Male	California	Human Resources	Baumber	Dodge

10	493	Male	Maryland	Product Management	Armatidge	Infiniti
11	492	Female	New York	Training	Minthorpe	Mercedes-Benz
12	491	Male	Nevada	Training	McIlmorow	Mercury
13	491	Female	New Mexico	Engineering	Blackborn	Mercedes-Benz
14	490	Female	Kentucky	Engineering	Heinert	Mercedes-Benz
15	489	Female	Florida	Marketing	Fraser	BMW
16	489	Male	California	Training	Menci	Audi
17	489	Male	Arizona	Engineering	Blythin	Saab
18	488	Female	California	Research and Development	Turford	Chevrolet
19	487	Female	Texas	Marketing	Forrington	Lexus
20	487	Female	Utah	Accounting	Deeks	BMW

Top 20 donors without University Education

Number	Donation	gender	State	job_field	university	car
1	493	Male	Louisiana	Business Development	Null	Chevrolet
2	492	Female	Florida	Business Development	Null	Ford
3	488	Male	Illinois	Product Management	Null	null
4	484	Male	Arkansas	Sales	Null	Ford
5	483	Female	Pennsylvania	Legal	Null	Mazda
6	482	Female	Virginia	Research and Development	Null	Hummer
7	480	Male	Tennessee	Human Resources	Null	Mercedes-Benz
8	480	Female	Arkansas	Engineering	Null	Nissan
9	476	Female	Tennessee	Product Management	Null	Rolls-Royce
10	473	Female	Indiana	Training	Null	Nissan
11	469	Male	Virginia	Research and Development	Null	Dodge

12	468	Female	Texas	Engineering	Null	null
13	468	Female	Michigan	Accounting	Null	Pontiac
14	467	Female	California	Business Development	Null	Nissan
15	467	Male	Mississippi	Legal	Null	null
16	463	Male	Missouri	Engineering	Null	Suzuki
17	462	Female	Pennsylvania	Support	Null	Mercedes-Benz
18	462	Male	Texas	Services	Null	Buick
19	461	Female	Iowa	Research and Development	Null	null
20	458	Male	Kentucky	Marketing	Null	null

As we can see, since the donors have 939 cars, it doesn't have any noticeable relationship with the number of donations. Critically looking at the result provided above, those with university education tend to donate big amount of donation compares with donors without university education. Another factor is male donates much higher than female. So, the target market is the male and female with university education.

CONCLUSION

It may be helpful to recall that the primary objective of this project is to understand the business problem of EDUCATION for All in order to increase donors in our database, increase the value and income of the donations from the donors. Having analyzed the data set given which are EFO_Donated_data and EFO_Donor_data, I found out our donors in our dataset lives in different states, work in different industries, our dataset have donors university education than other donors without university. Most of the donors have car leaving few donors without car and this can be ignored because there's no any correlation between having car and donating.

I also found out that the amount of donation of different donors vary with the maximum donation being \$500 and the minimum donation being \$5. But this also vary according to the frequency of the donation. Most donors donate just once or yearly rather than weekly or monthly. This caused a slow growth in increasing incomes from donation. So we should be ensure that we engage donors and remind them to donate as least monthly or quarterly. We can also convince our donors to have a standing order of donation in their bank account so as to not forget donating. Another thing to take note is that there are some donors that are not donating at all. From my observation, we have 753 donors that donates leaving 247 donors that have not donated. We should try to contact them and ask why they are not donating. Ensuring all donors to donates with improve the income of EDUCATION for All.

Another observation is that we should focus about getting more donors from other states apart from California, Texas and Florida since most donors that donated are from there.

Also another crucial point is that donors that are in the job field of Human Resources, Research Development, Project Management, and Business Development tend to donates more than other job field. This may due to the income they receives. So we can also target about getting more donors from those job field to increase the value of donation.

Lastly and more importantly, donors with university education are more likely to donate big than donors without university education. This does not mean that donors without university education doesn't donate big but compared to donors with university education they are not as many as those with university education.

In conclusion, despite the information given, we need to add more valuable information like contact such as phone and email of each donors, the economic trend of each state and average

income of each job field. Information such as movie genre, the language they speak, shirt size, and favorite colour are not useful in solving our business problem.

We need to be consistent in promoting our charity organization. Contacting our donors with university education to inform other about our charity organization is a good way. Also spreading our organization objectives across job field like human resources, project management, research development and business development can also be useful.