DATA ANALYST: SQL PORTFOLIO

PREPARED BY

GLORIA OLANIYAN

Professional Background

I graduated from Ladoke Akintola University of Technology, Ogbomosho, Oyo state, Nigeria, with a Second -Class upper CGPA of 3.9/5.0 B.Tech (honour) degree in Animal Production and Health. My first work experience was at the Ministry of Agriculture, Osun State Agricultural Development Cooperation OSADEC, Osogbo, Nigeria, where I acquired industrial training in my field of course. I am a member of the Nigeria Institute of Animal Science (NIAS) and a Graduate of Animal Scientist (GAS).

I served as an awareness manager in a non-governmental Organization, The Haemophilia Foundation of Nigeria (HFN) working under the World Federation of Haemophilia (WFH) in advocating, supporting and treating people with a bleeding disorder. I've worked as an Investment manager at Oxford Group of Companies, Lagos, Nigeria where I gave financial advice, trained my associates on prospecting, and led presentations in corporate firms. I'm skilled in managing people and building teams.

My affection for problem-solving, numbers and analytical thinking bring about my interest in Tech. In 2021, I started my journey in Tech through some pieces of training, I got a chance to attend a two-month Bootcamp for data analytics which included topics such as Data cleaning, Data Visualization, Data Extraction and Data Modelling. I have technical skills in Microsoft Excel, Power BI, Tableau, SQL and R.

Portfolio Outline

Introduction

I was in a situation given a position as Data Analyst working for the Charity, **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 donor frequency of our donors.
- Increase the value of donations in our database.

In two weeks, I needed to present the insights from the donation data to inform our fundraising strategy to increase donations in the following year.

I used the datasets EFO_Donation_Data and EFO_Donor_Data to answer the business problem.

EFO_Donation_Data contains the following data;- Donor ID, Donor first name, Donor last name, Donor Email address, Donor gender, Donor job field, Donation amount, Donor State of residence(US), Donor T-shirt size.

EFO_Donor_Data contains the following Data;- Donor ID, Frequency of donation, Donor University attended, Donor Car make, Donor second language, Donor Favourite colour, and Donor Favourite Movie genre.

I used the following SQL Queries to analyze the data, JOIN, ORDER BY, WHERE, BETWEEN, AND, OR, SUM(), COUNT(), AVG(), GROUP BY, HAVING.
I used Tableau to visualize the data and understood the problem by asking the right questions using the root cause analysis.

I got an insight from the report on the cause for the reduced donations in various states. I concluded and prepared the report for my team.

Root Cause Analysis

5 Why's Approach

To get deeper into the business problem, the why's approach was introduced.

During the course of the analysis, I discovered the number of donors was very low and the performance of the donation was very poor. I applied the root cause analysis to ask;-

- Why is the performance of donation low?

...because there were no or fewer donors in some states compared to others.

- Why do we have fewer donors in some states?

...because the donation method used in raising funds is not yielding.

- Why is the method not yielding?

...because we do not have enough useful information about the donors to know why they choose the organization and how they learned about us to use the approach to promote the organization.

- Why do we not have useful information about the donors?

...because we do not collect that useful information in the data.

- Why don't we collect useful information?

...because the strategy for raising funds is poor.

Insights

SQLite Database Management was used to find out the main insights. I was provided with EFO_Donation_Data and EFO_Donor_Data to answer the business problem.
The two datasets were imported into SQLite.

SELECT Function was used to select all the data from the tables

```
SELECT * FROM Donation_Data
SELECT * FROM Donor_Data2
```

To know the number of donors with **COUNT()**

```
SELECT COUNT(donation)
FROM Donation_Data;
```

To find the lowest value of donations with **Min()**

```
SELECT min(donation)
FROM Donation_Data;
```

Max() finds the largest value of the donations

```
SELECT max(donation)
FROM Donation_Data;
```

Sum() function sums up the total value of donations

```
SELECT sum(donation)
FROM Donation_Data
```

JOIN () function helps to combine rows from the two tables together using **ON ()** to identify the unique values of the two tables. **ORDER BY** helps in sorting numerically.

```
SELECT first_name, last_name, gender, job_field,donation,
state, Donor_Data2.donation_frequency
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
ORDER BY donation DESC
```

Top 20 Donors using **WHERE** clause and **ORDER BY** to sort

```
SELECT first_name, last_name, gender, job_field,donation,
state, Donor_Data2.donation_frequency, Donor_Data2.university
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE donation > 400
ORDER BY donation DESC
LIMIT 20;
```

GROUP BY functions to help in grouping gender and the number of donors

```
35 SELECT gender, COUNT(*)
36 FROM Donation_Data
37 JOIN Donor_Data2
38 ON Donation_Data.id = Donor_Data2.id
39 GROUP BY gender;
```

GROUP BY functions to help in grouping the total sum of donations by Gender

```
41 SELECT gender, sum(donation)
42 FROM Donation_Data
43 JOIN Donor_Data2
44 ON Donation_Data.id = Donor_Data2.id
45 GROUP BY gender;
46
```

Top 10 States by the sum of donations and number of donors

```
47 SELECT state, COUNT(*), sum(donation)
48 FROM Donation_Data
49 JOIN Donor_Data2
50 ON Donation_Data.id = Donor_Data2.id
51 GROUP BY state
52 ORDER BY sum(donation) DESC;
53 LIMIT 10;
```

Sum of Donation and number of donors by Donors' Job field

```
SELECT job_field, COUNT(*), sum(donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
GROUP BY job_field
ORDER BY sum(donation) DESC;
LIMIT 10;
```

Sum of Donation and number of donors by Donation Frequency

```
63 SELECT Donor_Data2.donation_frequency, COUNT(*), sum(donation)
64 FROM Donation_Data
65 JOIN Donor_Data2
66 ON Donation_Data.id = Donor_Data2.id
67 GROUP BY Donor_Data2.donation_frequency
68 ORDER BY sum(donation) DESC;
```

Top 10 donors by job field sorting by using the **HAVING** function

```sql
SELECT job_field, COUNT(*), sum(donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
GROUP BY job_field
HAVING COUNT(*) >50
ORDER BY sum(donation) DESC
LIMIT 10;
```

The sum of donations and count of donors by car make of donors

```sql
SELECT Donor_Data2.car, COUNT(*), sum(donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
GROUP BY Donor_Data2.car
ORDER BY sum(donation) DESC;
```

Top 10 male donors

```sql
SELECT first_name, last_name, gender, job_field, donation,
state, Donor_Data2.donation_frequency, Donor_Data2.university
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE donation > 400
AND gender = 'Male'
ORDER BY donation DESC
LIMIT 10;
```

Top 20 female donors

```sql
SELECT first_name, last_name, gender, job_field, donation,
state, Donor_Data2.donation_frequency, Donor_Data2.university
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE donation > 400
AND gender = 'Female'
ORDER BY donation DESC
LIMIT 20;
```

Top donors with and without university degree

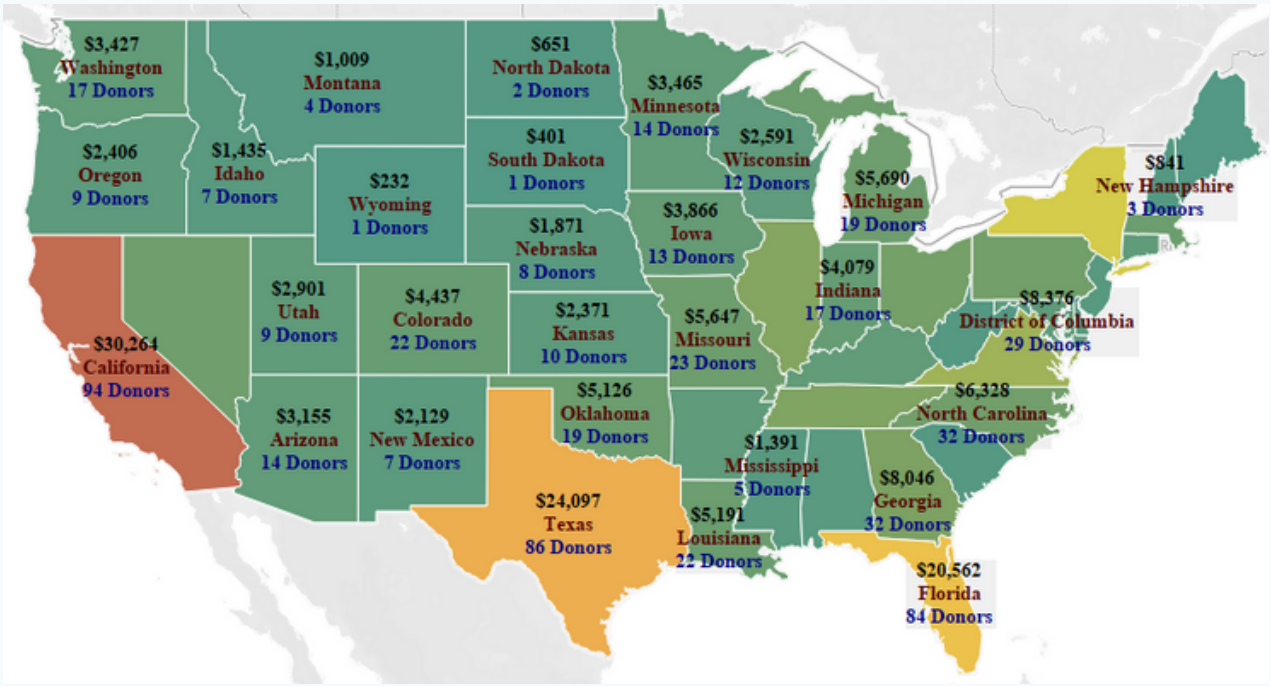
```sql
SELECT Donation_Data.donation, Donation_Data.gender, Donation_Data.state , Donor_Data2.donation_frequency, Donor_Data2.university
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE Donor_Data2.university IS NOT NULL
AND Donation BETWEEN 300 AND 500
ORDER BY donation DESC;

SELECT Donation_Data.donation, Donation_Data.gender, Donation_Data.state , Donor_Data2.donation_frequency, Donor_Data2.university
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE Donor_Data2.university IS NULL
ORDER BY donation DESC;
```

Tableau was used for data analysis and visualization. With Tableau, I was able to visualize the business problem.

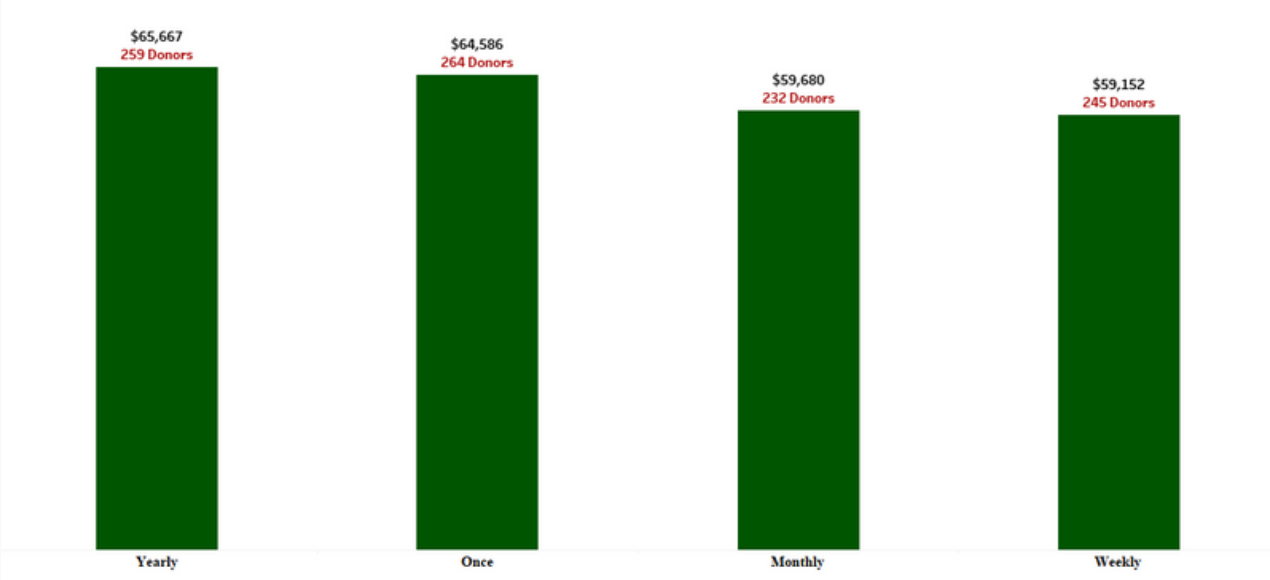
The table below shows the donors' job type by their donations and the number of donors. Donors with their job type as Human Resources had the highest donation and yet not the highest donors.

This map visualization shows the donors' state by their donations and the number of donors. California made the highest donation with the highest donors.

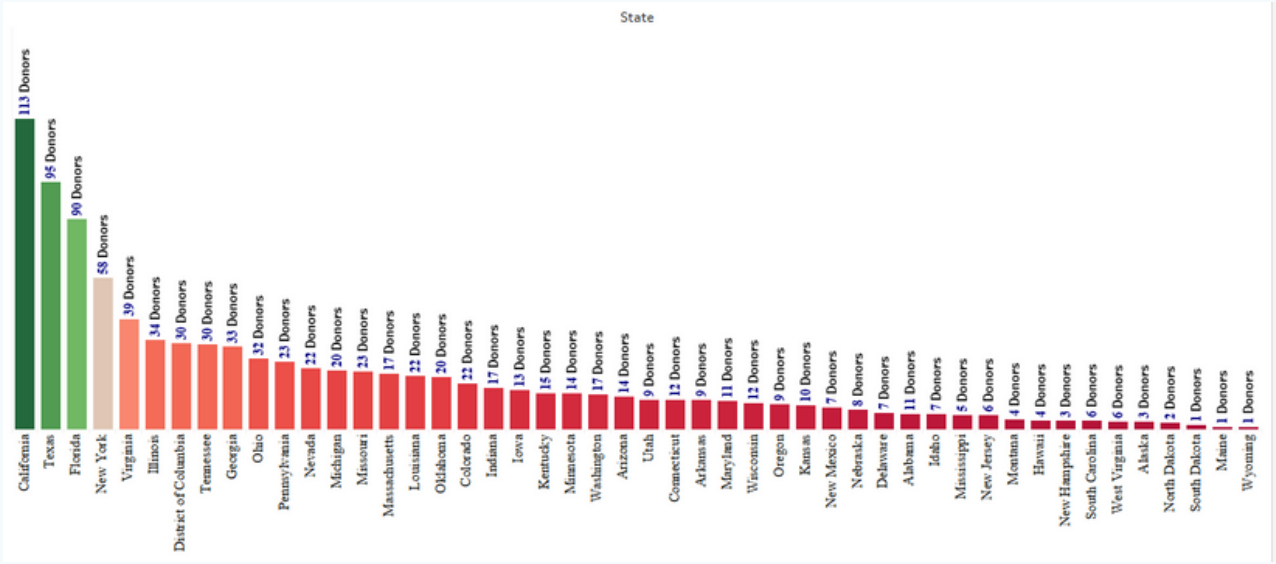


The Figure below shows the donors' gender by their donations and the donors' rate. The males donated more than females, whereas the female donors are more than male donors.

The below visualization shows the Donation frequency by the donation rates and the number of donors. Yearly donors donated more than others and once donors had the highest donors.



The visualization below shows the states by the donation rates and the number of donors. The top 3 states with the highest donation have the green colour.



Findings and Recommendations

From the analysis, I found out that :-

The sum of Donations received is $249,085
The number of Donors is 1000
The largest amount donated is $500
The lowest amount donated is $5
Donors are in 49 states in the United States.

FINDINGS 1: The top 10 male donors. Many of the top donors have university degree and more of them are from California.

first_name	gender	donation	state	donation_frequency	university
Beverlie	Male	500	Michigan	Yearly	Walasik
Wallie	Male	500	New York	Monthly	Leithgoe
Worthy	Male	498	Wisconsin	Monthly	Trotton
Amalea	Male	497	New York	Weekly	Rockcliffe
Tonnie	Male	494	California	Weekly	Cominetti
Nathaniel	Male	494	California	Monthly	Baumber
Beverlee	Male	493	Maryland	Monthly	Armatidge
Corbin	Male	493	Louisiana	Yearly	NULL
Emmit	Male	491	Nevada	Once	McIlmorow
Ludvig	Male	489	California	Monthly	Menci

FINDING 2: The findings show that males donated more than females. Female donors are more than male donors.

gender	count(*)	sum(donation)
Female	508	121457
Male	492	127628

FINDING 3: Top 10 Female Donors. Majority of the top donors have University degree and can also find many of them in California

first_name	gender	donation	state	donation_frequency	university
Clevie	Female	499	Virginia	Yearly	Sparhawk
Peder	Female	499	Delaware	Yearly	Antoszewski
Corbett	Female	494	California	Monthly	Coates
Hurley	Female	492	Florida	Weekly	NULL
Eddi	Female	492	New York	Yearly	Minthorpe
Babbette	Female	491	New Mexico	Monthly	Blackborn
Karilynn	Female	490	Kentucky	Monthly	Heinert
Charlotta	Female	489	Florida	Weekly	Fraser
Maura	Female	488	California	Weekly	Turford
Karena	Female	487	Texas	Weekly	Forrington

FINDINGS 4: 10 least Female Donors

first_name	gender	donation	state	donation_frequency	university
Sari	Female	14	California	Yearly	NULL
Avrom	Female	13	North Carolina	Weekly	Tapenden
Ferdinanda	Female	10	Florida	Weekly	NULL
Marys	Female	9	Tennessee	Once	Boichat
Peta	Female	9	Idaho	Once	Goodlatt
Mathilde	Female	9	Oklahoma	Once	Simoncelli
Matthiew	Female	7	Texas	Yearly	Gurr
Marline	Female	7	Florida	Monthly	NULL
Korie	Female	7	Ohio	Yearly	Traher
Garrek	Female	6	Missouri	Yearly	Seeborne

FINDINGS 5: We have many donors who donated yearly and once and more donations were made on yearly basis

donation_frequency	COUNT(*)	sum(donation)
Yearly	259	65667
Once	264	64586
Monthly	232	59680
Weekly	245	59152

FINDINGS 6: Top states with the highest donations. California, Texas and Florida are the top 3 States that have many donors donated to our organization.

state	count(*)	sum(donation)
California	113	30264
Texas	95	24097
Florida	90	20562
New York	58	14759
Virginia	39	10750
Illinois	34	8674
District of Columbia	30	8376
Tennessee	30	8316
Georgia	33	8046
Ohio	32	6876
Pennsylvania	23	6574
North Carolina	33	6328

FINDING 7: The least states with the lowest donors. States that have less than 10 donors need to be worked on to improve the donation rates in the subsequent years.

state	count(*)	sum(donation)
Kansas	10	2371
New Mexico	7	2129
Nebraska	8	1871
Delaware	7	1569
Alabama	11	1446
Idaho	7	1435
Mississippi	5	1391
New Jersey	6	1376
Montana	4	1009
Hawaii	4	875
New Hampshire	3	841
South Carolina	6	819
West Virginia	6	793
Alaska	3	734
North Dakota	2	651
South Dakota	1	401
Maine	1	258
Wyoming	1	232

FINDINGS 8: The highest 10 donors with University degrees

donation	gender	state	university
500	Male	Michigan	Walasik
500	Male	New York	Leithgoe
499	Female	Virginia	Sparhawk
499	Female	Delaware	Antoszewski
498	Male	Wisconsin	Trotton
497	Male	New York	Rockcliffe
494	Male	California	Cominetti
494	Female	California	Coates
494	Male	California	Baumber
493	Male	Maryland	Armatidge

FINDING 9: The least 10 donors with University Degree

donation	gender	state	university
5	Male	Oklahoma	Eyrl
6	Male	California	Marfell
6	Female	Missouri	Seeborne
6	Female	California	Kleint
6	Male	North Carolina	Franschini
7	Female	Texas	Gurr
7	Female	Ohio	Traher
9	Female	Tennessee	Boichat
9	Female	Idaho	Goodlatt
9	Female	Oklahoma	Simoncelli

FINDINGS 10: Top 10 donors without University degree

donation	gender	state	university
493	Male	Louisiana	NULL
492	Female	Florida	NULL
488	Male	Illinois	NULL
484	Male	Arkansas	NULL
483	Female	Pennsylvania	NULL
482	Female	Virginia	NULL
480	Male	Tennessee	NULL
480	Female	Arkansas	NULL
476	Female	Tennessee	NULL
473	Female	Indiana	NULL

FINDING 11: The least 10 donors without a university degree.

donation	gender	state	university
5	Female	Alabama	NULL
6	Male	Colorado	NULL
7	Female	Florida	NULL
8	Male	Florida	NULL
9	Male	Florida	NULL
10	Female	Florida	NULL
13	Male	Arkansas	NULL
14	Female	California	NULL
17	Female	New York	NULL
19	Male	Missouri	NULL

Recommendations

After analyzing EFO_Donation_Data and EFO_Donor_Data to understand the business problem, I discovered that many states need to be worked on by making a strong awareness to support our charity organisation by donating. Especially the states where we have less than ten donors.
Platforms, where donations are made, should be improved and more platforms should be created and monitored to know where there is progress and where there is not to increase the donation rates.

Data like Donor Car make, Donor second language, Donor Favourite colour, and Donor Favourite Movie genre should be excluded and more useful information like how they know about the organization, Nationality, and age should be included,
Males tend to donate more than females even though we have females donor than males.

To increase the number of donors, Chapter meetings and campaigns for awareness should be organized monthly or annually across the 49 states where donations are made. The higher the donor, the higher the donation. More concentration should be given to the elites in these populated states like California, Texas and Florida to boost the donation rates.

Conclusion

When we have more channels for promoting the charity, the platforms are monitored, campaigns are organized in all the states where we have donors, useful information about the subscribers is collected and concentrations are given to states with low performance, this will, in turn, improve the organization and there will be progress in donations which will increase the donation rates,
there will be an increase in donations and
we will have more donors in the organization.