

DATA ANALYST: SQL PORTFOLIO

PREPARED BY

Ephraim Michael



Professional Background

Educational Background:

- M.Eng in Telecommunications Engineering from the Federal University of Technology Minna 2017-2020
- B.Eng. in Electrical and Computer Engineering from Federal University of Technology Minna 2004-2009

Professional Background:

- Assistant Chief Engineer at Scientific Equipment Development Institute Minna Nigeria from 2021 to date
- Senior Engineer at Scientific Equipment Development Institute Minna Nigeria from 2018 to 2020
- Engineer I at Scientific Equipment Development Institute Minna Nigeria from 2015 to 2017
- Engineer II at Scientific Equipment Development Institute Minna Nigeria from 2012 to 2014.

Skills

- Data Analyst
 - Python, Excel, Tableau, SQL
- GUI developer
 - C#
- Embedded system design
 - PIC, AVR, Raspberry pi, Arduino, ESP
- Embedded system programming
 - Assembly, C++, Python,
- Machine Learning and Deep Learning
 - Python

Portfolio Outline

Professional Background	1
Table of Contents	2
Introduction	3
Root Cause Analysis	4
Insights	5
Findings and Recommendations	6
Conclusion	7





Introduction

Situation:

The charity organization called "Education for All" intends to hold a fundraising program to support education. The goal of the fundraising is to increase the number of donors, their donation frequency and to increase the amount donated. Prior to the event, all stake holders in the organization are to discourse strategies that will help achieve these goals

Task:

As a data analyst, I was to make suggestions based on insights from data provided from an SQL data base.

Action:

To achieve results, root cause analysis was carried out. Afterwards, SQL query was used to derive insights. these insights was represented visually using tableau

Result:

At the end of the analysis, characteristics of donors that could help achieve the set goals was identified..

Root Cause Analysis

- Root cause analysis is a technique invented by Sakichi Toyoda that involves asking 5 WHYs to be able to unveil deep insight about a business problem.. Note that it is not compulsory for the WHYs to be 5.. However, before the fifth why is exhausted, the root cause will be discovered.

THE Root Cause Analysis

- Why did we not have high donations, frequency of donation and number of donors
 - Ans: we put equal concentration on states considered while seeking funds.
- Why did we put equal concentration on states considered?
 - Ans: We had no information of which state and the characteristics of the people in the state that could donate more and assumed that more donations could come from larger state
- Why do we not have information on these characteristics so we could target them?
 - Ans: We don't know how characteristics such as gender, job field,, ability to own a car, and donation frequency relates with the amount of donations that can be achieve in each state.

Insights

Question 1

List 10 state in descending order where the highest sum of donations will be gathered.

Question 2

Does the size of the state correlates with the sum of donation

```
Select state, sum(donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id =
Donor_Data2.id
GROUP by state
ORDER by sum(donation) DESC;
```

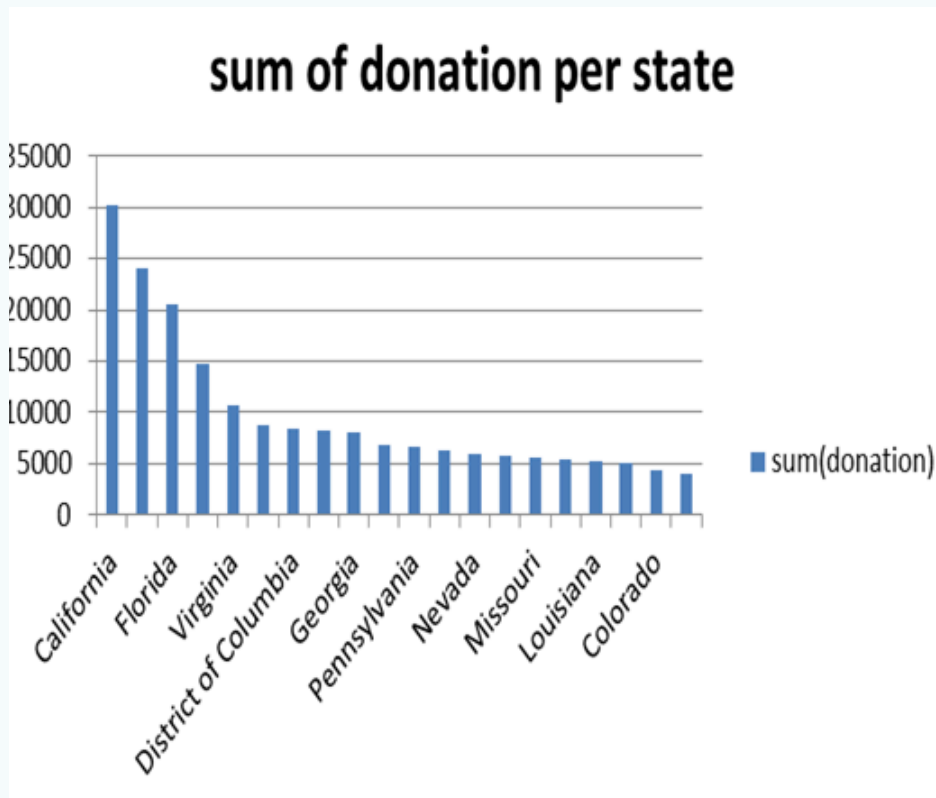
Fig 1 shows the SQL query was used to join the two data base. The sum donations in each state was investigated. Fig 2 shows the first 10 highest donation displayed in excel sheet.

Fig 1. SQL query for sum of donation per state

1	state	gender	sum(donation)
2	California	Male	16836
3	Texas	Male	13687
4	Florida	Male	10097
5	New York	Male	7408
6	Virginia	Male	6298
7	Ohio	Male	4810
8	Nevada	Male	4206
9	Tennessee	Male	4154
10	District of	Male	4054
11	Illinois	Male	4015

Fig. 2 Excel result of the SQL query in Fig 1

Insights



The graph in Fig 3 shows the 10 to countries in order of the sum of their donation. Fig 4 shows that the lowest sum of donation is \$232 and the highest sum of donation is \$30624 made by California. Also, it is observed that the size of the state does not determine the amount donated.

Fig 3 graphical illustration of the 10 highest sum of donation per state

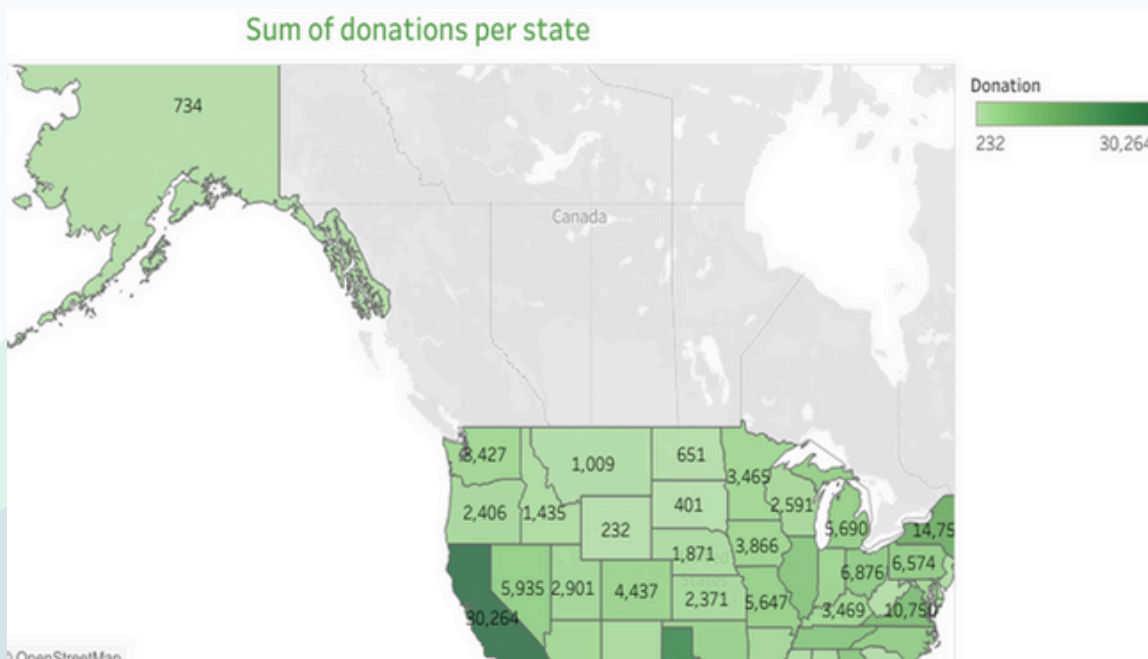


Fig 4 Map representation of the states with sum of donation made

Insights

Question 3

List 10 state with the highest average donation.

Question 4

Is the average donation from the state correlating with the sum in each state?

```
Select state, ROUND (AVG (donation), 2)

FROM Donation_Data

JOIN Donor_Data2

ON Donation_Data.id = Donor_Data2.id

GROUP by state

ORDER by ROUND (AVG (donation), 2) DESC

LIMIT 10;
```

Fig 5. SQL query to get the average donation per state

1	state	ROUND (AVG(donation),2)		
2	South Dakota	401		
3	North Dakota	325.5		
4	Utah	322.33		
5	Massachusetts	316.76		
6	Arkansas	315.22		
7	New Mexico	304.14		
8	Iowa	297.38		
9	Pennsylvania	285.83		
10	Michigan	284.5		
11	New Hampshire	280.33		

Fig 6 Result of the query

Insights

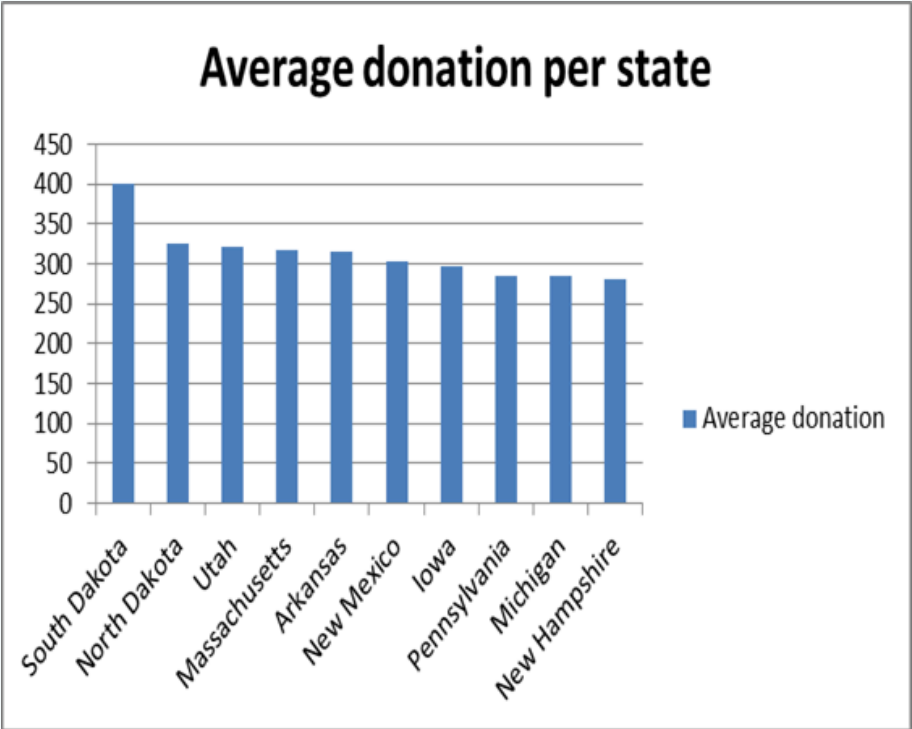


Fig 7 graphical illustration of the 10 highest average donation per state

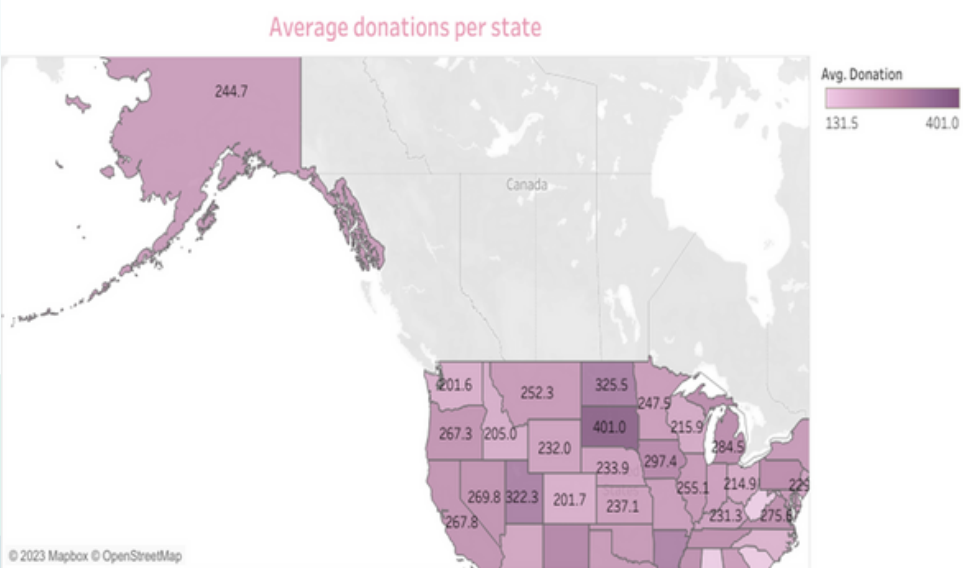


Fig 8 Map representation of the states with average of donation

Fig 7 shows the graphical representation of the queried average donations per state. fig 8 show the map representation of this state. from the figure it shows that the minimum average donation was \$131.5 and the maximum average donation was \$401.5. Also it was observed that these 10 state with highest average donations are not the same as those with highest sum donation in fig 3

Insights

Question 5

List from the top ten states, which gender has the higher sum of donation per state

```
SELECT state, gender, SUM (donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE gender ='Male'
GROUP by state
ORDER by SUM (donation) DESC
LIMIT 10;
```

Fig 9. SQL query to get the sum donation per state from male

1	state	gender	sum(donation)	
2	California	Male	16836	
3	Texas	Male	13687	
4	Florida	Male	10097	
5	New York	Male	7408	
6	Virginia	Male	6298	
7	Ohio	Male	4810	
8	Nevada	Male	4206	
9	Tennessee	Male	4154	
10	District of	Male	4054	
11	Illinois	Male	4015	

Fig 10 Result of the query

Insights

```
SELECT state, gender, SUM (donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE gender ='Male'
GROUP by state
ORDER by SUM (donation) DESC
LIMIT 10;
```

Fig 11. SQL query to get the sum donation per state from male

1	state	gender	SUM(donation)	
2	California	Female	13428	
3	Florida	Female	10465	
4	Texas	Female	10410	
5	New York	Female	7351	
6	Illinois	Female	4659	
7	Virginia	Female	4452	
8	District of	Female	4322	
9	Georgia	Female	4246	
10	Tennessee	Female	4162	
11	Massachu	Female	4090	

Fig 12 Result of the query

Insights

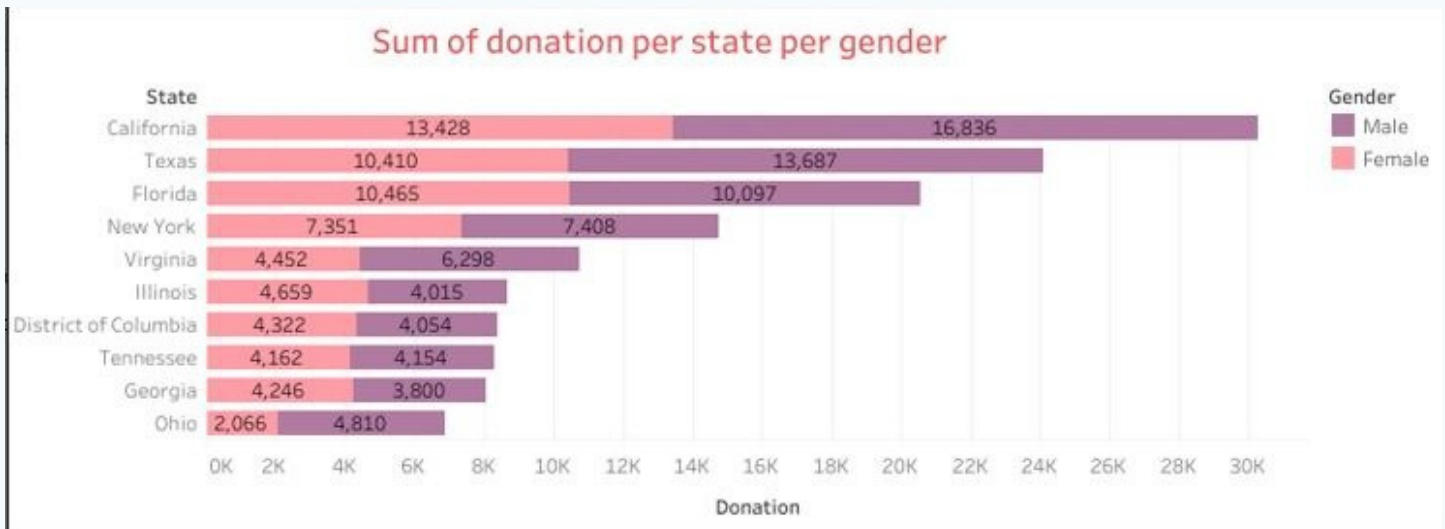


Fig 13 graphical illustration of the 10 highest average donation per state

The graphical representation in Fig 13 shows that, the dominant donors in California, Texas, New York, Virginia and Ohio are the men while others are female. this means the gender matters in each state.

Question 6

Which job does the highest sum of donation

```
SELECT job_field, SUM (donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
--where gender ='Male'
GROUP by job_field
ORDER by SUM (donation) DESC;
```

Fig 14. SQL query to get the sum donation per job field

Insights

Donation per job field

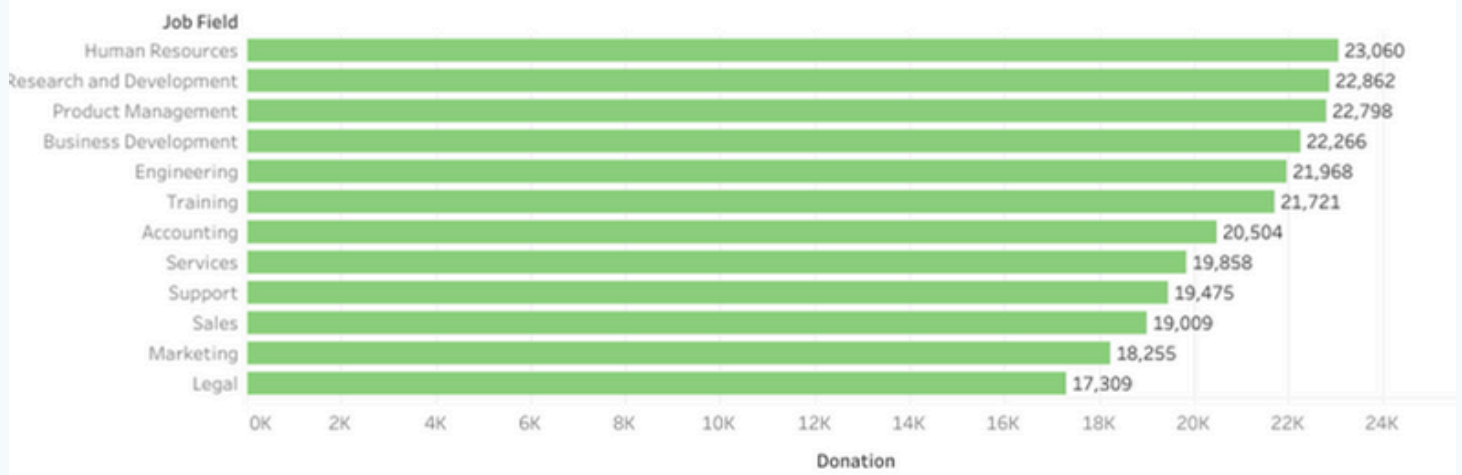


Fig 15 graphical illustration of the highest sum donation per job field

Question 7

Which state does the highest average donation at each defined donation frequency?

```
SELECT state, ROUND ( AVG (donation), 2)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE donation frequency='Once'
GROUP by state
ORDER by ROUND (AVG(donation),2) DESC
LIMIT 10;
```

Fig 16. SQL query to get the average donation per state done once

Insights

```
SELECT state, ROUND (AVG (donation), 2)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE donation frequency='Weekly'
GROUP by state
ORDER by ROUND (AVG (donation),2) DESC
LIMIT 10;
```

Fig 17. SQL query to get the average donation per state done weekly

```
SELECT state, ROUND (AVG (donation), 2)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE donation frequency='Monthly'
GROUP by state
ORDER by ROUND (AVG(donation),2) DESC
LIMIT 10;
```

Fig 18. SQL query to get the average donation per state done monthly

Insights

```
SELECT state, ROUND (AVG (donation), 2)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id
WHERE donation frequency='Yearly'
GROUP by state
ORDER by ROUND (AVG (donation),2) DESC
LIMIT 10;
```

Fig 19. SQL query to get the average donation per state done yearly

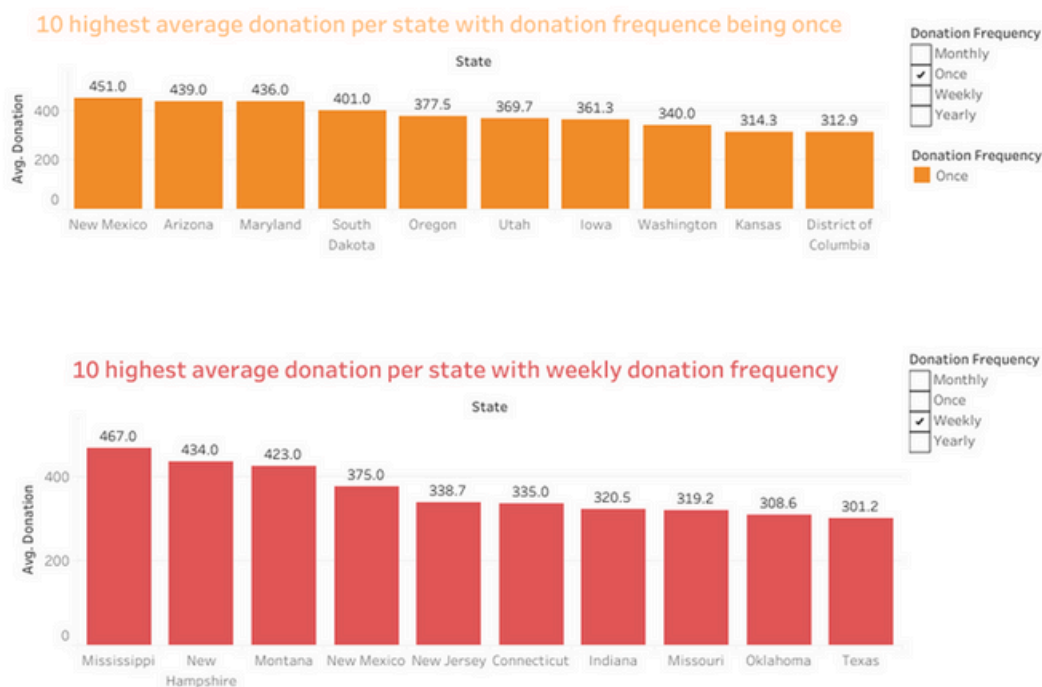


Fig 20. graphical illustration of average donation per state done once and weekly

Insights

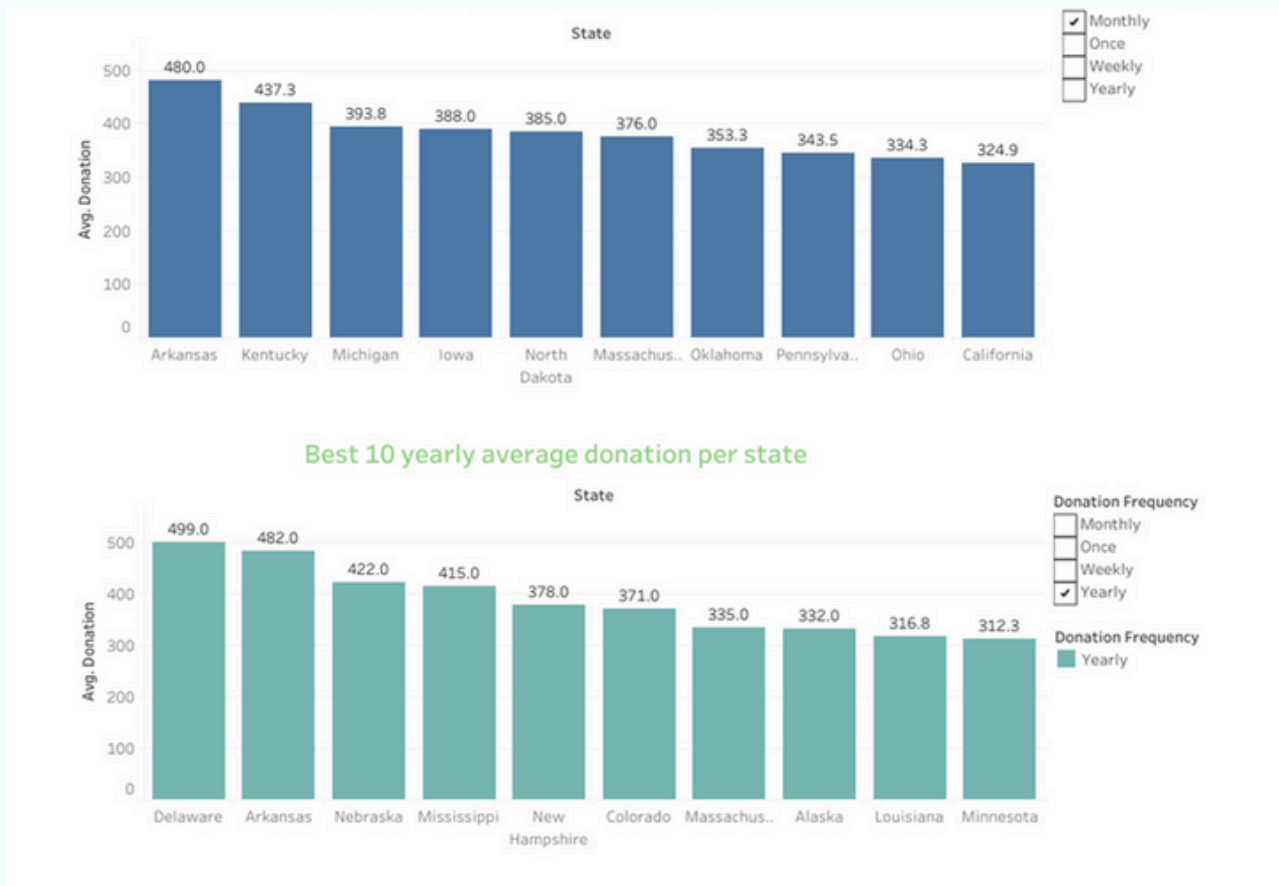


Fig 21. graphical illustration of average donation per state done monthly and yearly

Question 8

Does owning a car correlate to the amount donated?

Insights

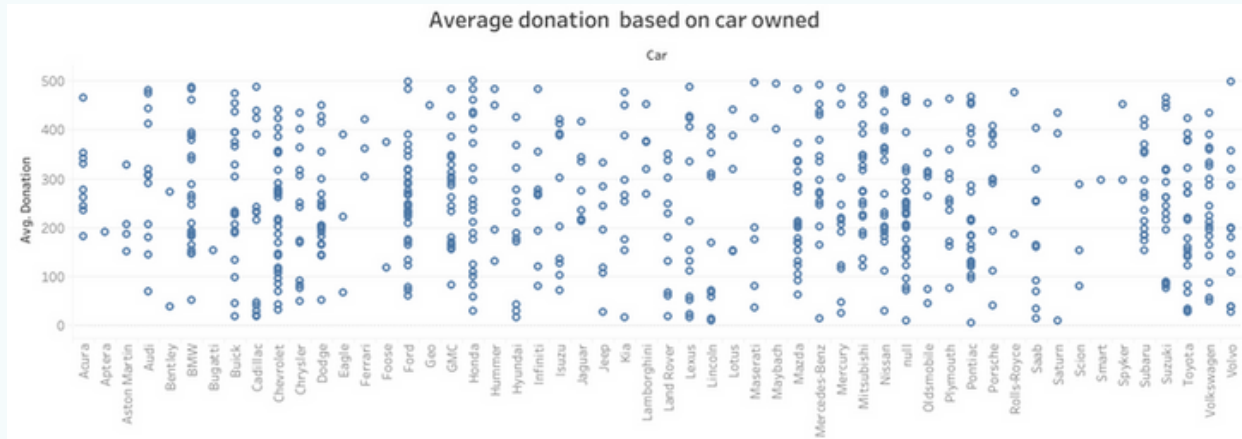


Fig 21 graphical illustration of average donation per car owner

This shows that owning a car does not determine the average amount donated.

Findings and Recommendations

To increase the number of donors and the amount donated, the following has to be considered seriously.

- * The intensity for canvassing for funds in states like California, Texas, Florida, New York, Virginia, Illinois, District of Colombia, Tennessee, Georgia and Ohio should be considered in the order of listing first before any other available state.

- * Among these states, male donors should be targeted seriously in California, Texas, New York, Virginia and Ohio as they show more potential to giving than Female donor. Also in Florida, Illinois, District of Colombia Tennessee and Georgia, female donors should be the target. In other words, female donors should be considered seriously compared to male donors.

- * During the pitching for funds, the jobs of potential donors should be considered preferentially as listed: Human Resource, Research and Development, Project Management, business development, Engineering Training , Account, services, Support, Sales, Marketing, Legal. this will increase the amount to be donated.

- * To increase the donation frequency, if the targets are states that will donate once is to be considered, then New Mexico, Arizona, Maryland , South Dakota, Oregon, Utah, Iowa, Washington. Kansas and District of Colombia should be the first 10 state to be considered in the order of listing.

- * If the targets are states that will donate weekly, then Mississippi, New Hampshire, Montana, New Mexico, New Jersey, Connecticut, Indiana, Missouri, Oklahoma and Texas should be the first 10 state to be considered in the order of listing.

- * If the targets are states that will donate Monthly, then Arkansas, Kentucky, Michigan, Iowa, North Dakota, Massachusetts, Oklahoma, Pennsylvania, Ohio and California should be the first 10 state to be considered in the order of listing.

- * If the targets are states that will donate yearly, then Delaware, Arkansas, Nebraska, Mississippi, New Hampshire, Colorado, Massachusetts, Alaska, Louisiana and Minnesota should be the first 10 state to be considered in the order of listing.

Conclusion

To raise funds for an inclusive education for a charity organization, the data of past fundraising events have been explored and insights gotten is used to proffer advise on strategies to be done to increase the amount of funds and the donation frequency.

The two data base was joined and queried using SQL. Tableau was used to aid data visualization.