DATA ANALYST 2: SQL PORTFOLIO

BY: NJOKU ESTHER UCHECHI

PROFESSIONAL BACKGRROUND

- As a graduate of Microbiology from **Imo State University**, Owerri, Imo State, Nigeria. My course work has provided me with an in-depth understanding of microbial physiology, pathogenesis, genetics, public health as well as statistical methods required for quantitative analysis. During my studies, I focused on analyzing the interactions between microorganisms and their environment and accessing the impact of microorganisms on human health.

- I was opportune to intern at the medical laboratory department of **Nigerian Navy Reference Hospital**, Ojo, Lagos, Nigeria. I had a first hand laboratory experience using a wide variety of laboratory techniques and also assisted in recording, analyzing and interpreting data from the laboratory test results which were used to prepare reports.

- I have been fascinated about the relationship of data analysis with the health sector and its importance in policy making and improvement of healthcare of individual and the population at large. Hence, my venture into data analysis class.

- I enrolled into Entrylevel's Data Analyst 1 as my first step in data analysis, where I learnt to clean data and create pivot tables and visualization using Microsoft Excel and Tableau. I worked on several projects where I cleaned data, created insights/visualizations and reports.

- Subsequently, I enrolled into Entrylevel's Data Analyst 2 involving SQL. I am a critical thinker and a good team player with good analytical skills, open to new innovations, long-term cooperation, new challenges and discoveries.

TABLE OF CONTENT

INTRODUCTION

- As a 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, where my objectives are to;
    1. Increase the number of donors on our database
    2. Increase the donation frequency of our donors
    3. Increase the value of donations in our database
- I needed to present my insights from the donation data provided and give information on my fund raising strategy to increase donations for the following year to my team.
- I applied the SQL commands: JOIN, ORDER BY, WHERE, BETWEEN, AND, OR, SUM(), COUNT(), AVG(), GROUP BY and HAVING.

BUSINESS PROBLEM

1. I figured out that there were low turnouts in the funds raised. That is, the funds raised for the charity was barely enough.

2. The data that should be collected are the donation data and the donor data to solve the business problem. Insights, analysis and data visualizations from the datasets were to be presented to my team in a fortnight.

QUESTIONS FOR BETTER UNDERSTANDING

1. How many donors do we have currently?
2. Who are the top 20 donors?
3. What is the total amount of donations?
4. What is the frequency of donations?
5. Are the amount of donations dependent on job field?
6. Which state has the highest amount of donors?
7. What amount of donations are received monthly?

ROOT CAUSE ANALYSIS

I applied root cause analysis using the "5 whys" technique to understand the cause of the business problem: **Low turnouts in the funds raised.**

1. **Q:** Why is the turnout low?

   A: The fundraising for charity events were not effectively promoted to the target audience.

2. **Q:** Why were the events not effectively promoted?

   A: There was lack of marketing and outreach strategy.

3. **Q:** Why was there lack of marketing and outreach strategy?

   A: The fundraising team does not have enough resources for event marketing.

4. **Q:** Why doesn't the fundraising team have enough resources?

   A: The budget allocation for event marketing is limited.

5. **Q:** Why is the budget allocation limited?

   A: The overall fundraising strategy has not prioritized event marketing.

INSIGHTS FROM THE ANALYSIS

- I was provided with Donation_Data and Donor_Data2 relational databases to answer the business problem and I used SQLite Online Database Management System to find my insights from the dataset.
- The Donation Dataset has the following data: Id, first_name, last_name, email, gender, job_field, donation state and shirt size.
- The Donor Dataset has the following data: Id, donation_frequency, university, car, second_language, favourite_colour, and movie_genre.
- I imported both datasets into SQLite and joined both tables using the SELECT, JOIN and ON commands:

```
SELECT *
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id = Donor_Data2.id ;
```

- To find the number of donors we have currently, I used the COUNT() function thus:

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

- To find the top 20 donors, I used the SELECT, FROM, ORDER BY, DESC and LIMIT function thus:

```
SELECT *
FROM Donation_Data
ORDER BY donation DESC
LIMIT 20 ;
```

- To find the total number of donations, I used the SELECT, SUM() and FROM functions thus:

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

- To find the frequency of donations done once, weekly, monthly and yearly, I used the SELECT, FROM, JOIN, ON and WHERE functions thus:

```
SELECT Donation_Data.first_name, Donation_Data.donation, Donor_Data2.donation_frequency
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id=Donor_Data2.id
WHERE donation_frequency = 'Once' ;
```

```
SELECT Donation_Data.first_name, Donation_Data.donation, Donor_Data2.donation_frequency
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id=Donor_Data2.id
WHERE donation_frequency = 'Weekly' ;
```

```
SELECT Donation_Data.first_name, Donation_Data.donation, Donor_Data2.donation_frequency
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id=Donor_Data2.id
WHERE donation_frequency = 'Monthly' ;
```

```sql
SELECT Donation_Data.first_name, Donation_Data.donation, Donor_Data2.donation_frequency
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id=Donor_Data2.id
WHERE donation_frequency = 'Yearly' ;
```

- To find the State with the highest amount of donors, I used the SELECT, COUNT(), FROM, ORDER BY and DESC functions thus:

```sql
SELECT state, COUNT (*)
FROM Donation_Data
GROUP BY state
ORDER BY COUNT (*) DESC;
```

- To find the amount of donations received once, monthly, weekly and yearly, I used the SELECT, FROM, JOIN, ON and WHERE functions thus:

```sql
SELECT Donation_Data.donation, Donor_Data2.donation_frequency, sum(donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id=Donor_Data2.id
WHERE donation_frequency= 'Yearly' ;
```

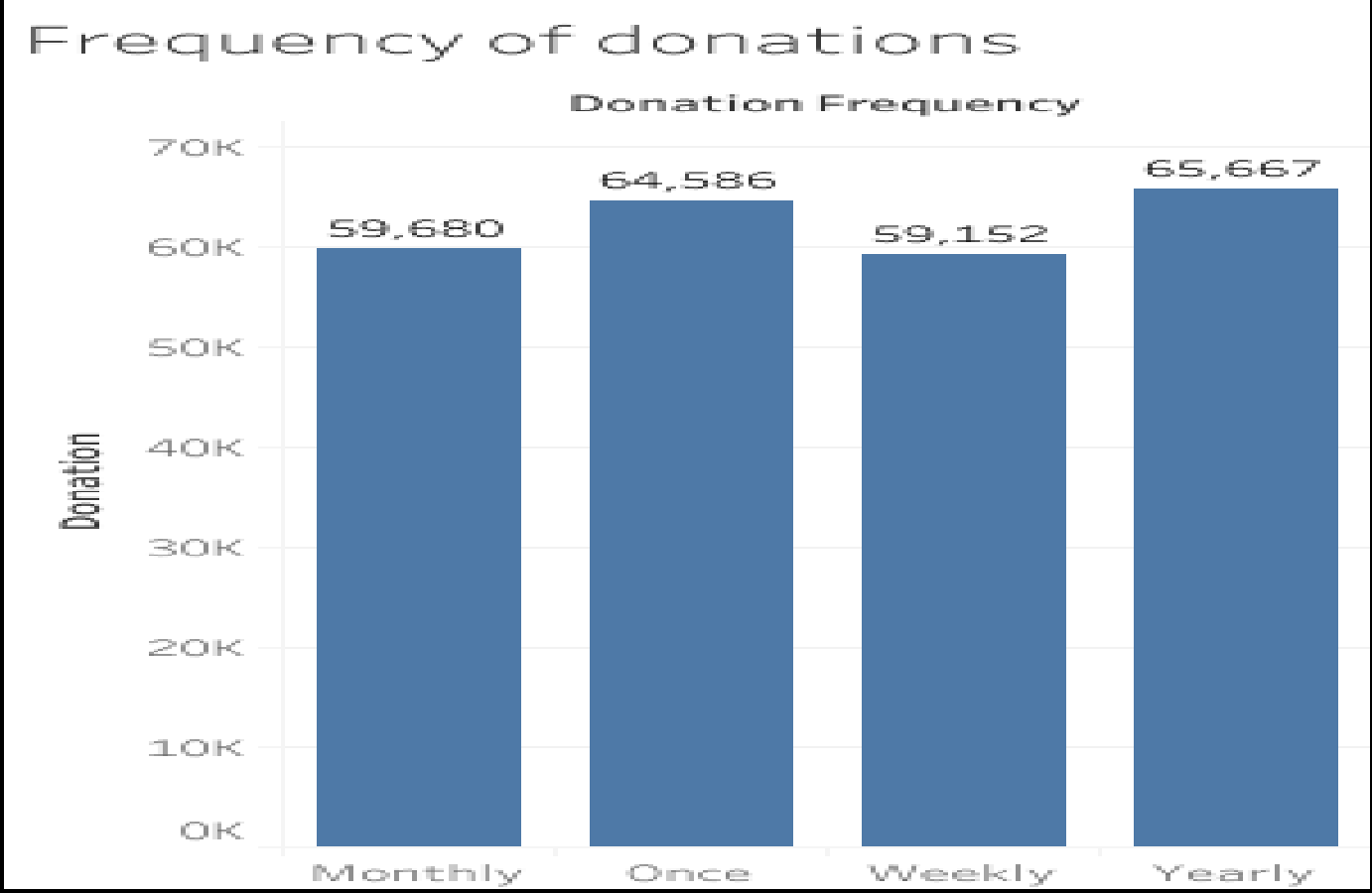
```sql
SELECT Donation_Data.donation, Donor_Data2.donation_frequency, sum(donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id=Donor_Data2.id
WHERE donation_frequency= 'Monthly' ;
```

```sql
SELECT Donation_Data.donation, Donor_Data2.donation_frequency, sum(donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id=Donor_Data2.id
WHERE donation_frequency= 'Weekly' ;
```

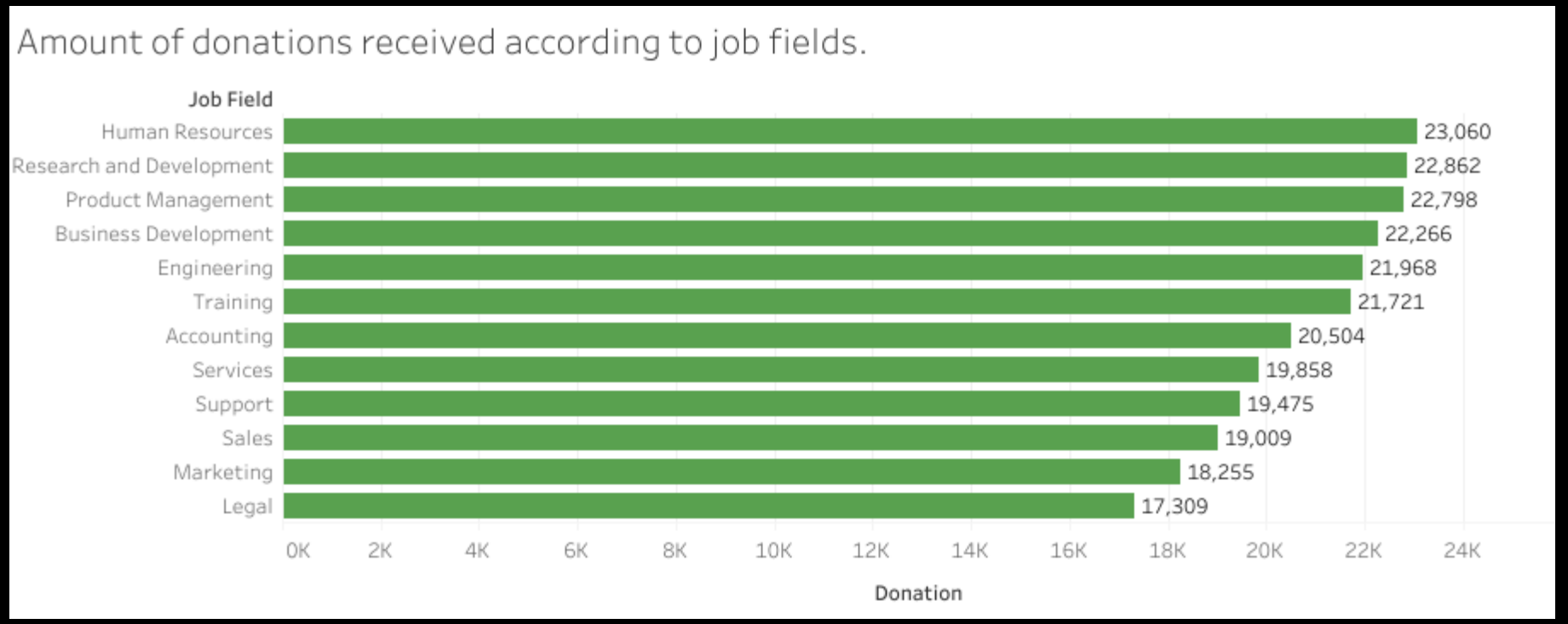
```sql
SELECT Donation_Data.donation, Donor_Data2.donation_frequency, sum(donation)
FROM Donation_Data
JOIN Donor_Data2
ON Donation_Data.id=Donor_Data2.id
WHERE donation_frequency= 'Once' ;
```

TABLEAU VISUALIZATIONS

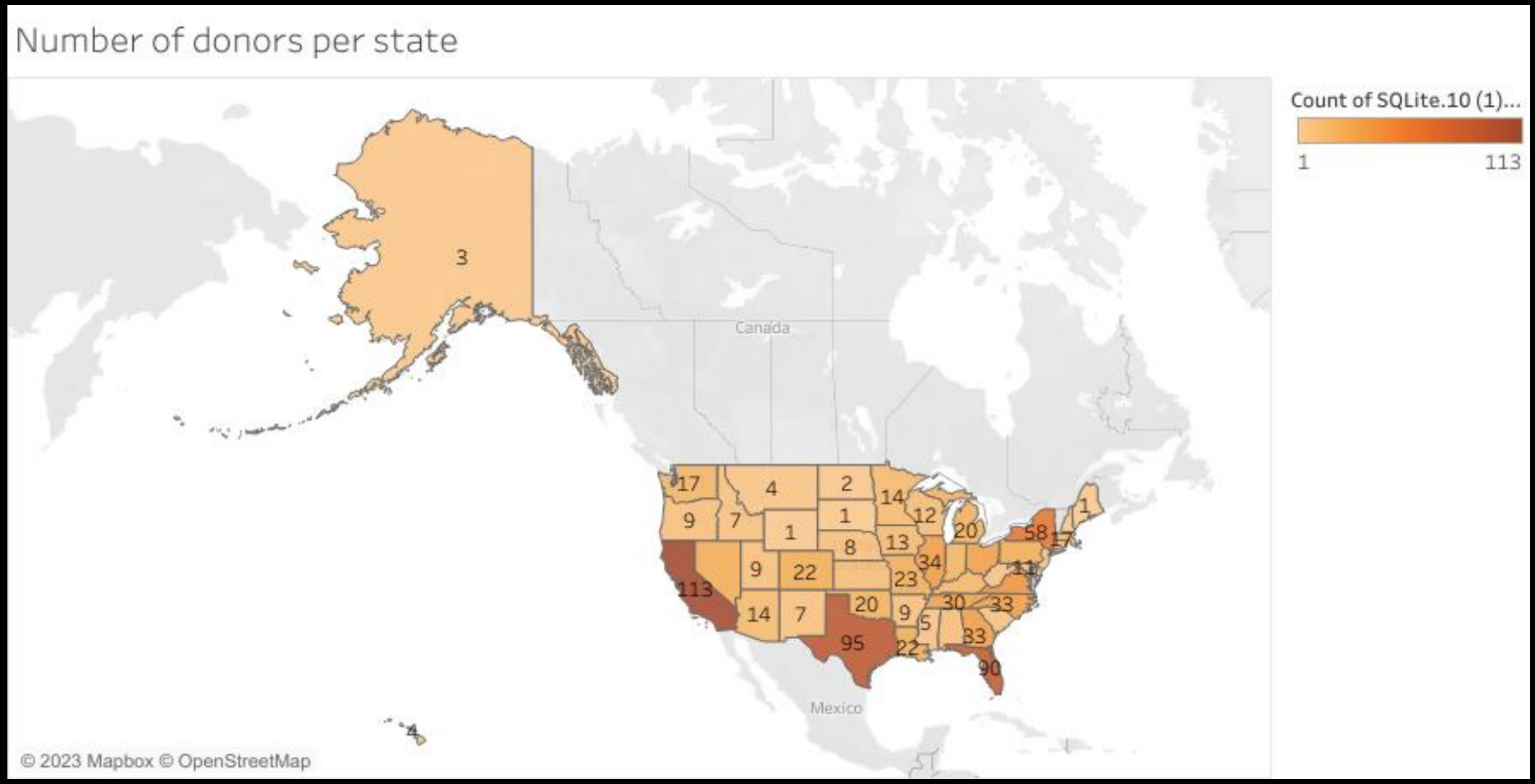
- The vertical bar chart represents the frequency and amount of donations received once, weekly, monthly and yearly respectively.



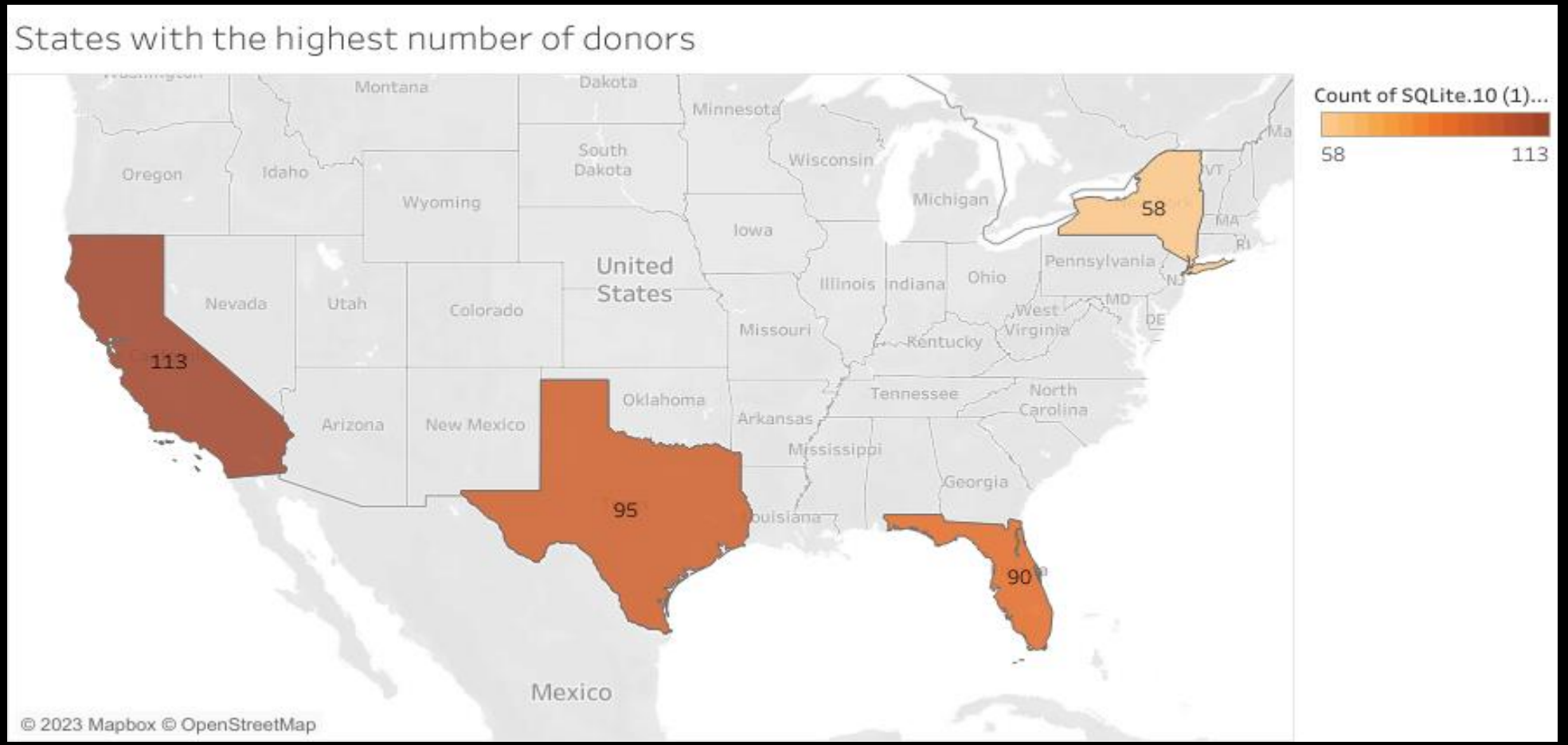
- The horizontal bar chart below shows the amount of donations received from each job field with Human resources having the highest amount of donations. Thus, the amount of donations are not dependent on the job field.



- The mapped chart below shows the states and the number of donors in each state with Maine, South Dakota and Wyoming having the lowest number of one donor each.



- The mapped chart below shows the state our highest donors come from, with California having the highest number of 113 donors, followed by Texas with a total number of 95 donors, Florida with a total number of 90 donors and New York having 58 donors respectively.



FINDINGS AND RECOMMENDATIONS

From my analysis, I found out that

- The total number of donors we have are 1000
- The total sum of donations collected are $249085
- The largest amount of donation collected is $500
- The smallest amount of donation collected is $5

From my insights, I figured out that out of the 1000 donors we have on our data base, the donations received are as follows:

DONATION_FREQUENCY	DONATION	SUM(DONATION)
Yearly	255	65667
Monthly	178	59680
Weekly	28	59152
Once	292	64586

- The table above indicates that we have 753 active donors who donate with different frequency. While the remaining 247 do not donate at all. Also, I noticed that most of our active donors donate mostly once and yearly. I suggest there should be an increase in the frequency of donations.

The top 10 states with the highest donors include:

STATE	COUNT(*)	SUM(DONATION)
California	113	30264
Texas	95	24097
Florida	90	20562
New York	58	14759
Virginia	39	10750
Illinois	34	8674
North Carolina	33	6328
Georgia	33	8046
Ohio	32	6876
Tennessee	30	8316

- From the table above, I figured out that the top 6 states with the highest number of donors has the highest amount of donations received. However, we can't fully come to a conclusion that the donations are dependent on the number of donors from each state. This is because the other states(North Carolina, Georgia, Ohio and Tennessee) are interchanged in the number of donors and donations. That is to say, some states with lesser number of donors have a higher amount of donations than others with slightly higher number of donors.

CONCLUSION

- Having analyzed the donation dataset and donor dataset to help education for all raise funds for charity and have a frequent donation, I found out that we have different donors, the rates at which they donate, different job fields and other factors.

- Although we have people who donate once, monthly, weekly and yearly basis, the outcome of our monthly and weekly donors is very low compared to our once and yearly donors. I suggest we work on the improvement of the frequency of donations. I also figured out that we have inactive donors on our database, we should try to reach out to them since we have their details.

- We also need to attract more donors from other states so there will be some sort of balance as there is a wide margin between donors from California, Texas and Florida than other states.

- We should also budget for marketing strategies and outreaches which would attract more donors and help increase the donations needed.