DATA ANALYST: SQL PORTFOLIO

PREPARED BY

Tolu Mayowa Adesina
sundaytolu4@gmail.com

Professional Background

I am a Data Analyst with experience in developing analytics solutions, maintaining data pipelines, delivering high level insights & recommendations, creating data-driven strategies that informs marketing initiatives geared towards optimizing ROI to drive product and business growth.

I have Demonstrated history of success in deploying analytics and experimentation to drive business & product growth in several firms. I have also led various end to end analytics initiatives across diverse business verticals & industries including Retail, education services and Tech.

I strongly believe firms will perform optimally through strategic positioning and strategic Analysis with Data. I am passionate about leveraging analytics & machine learning technologies in creating strategies geared towards optimizing marketing campaigns, improving product features, enhancing website revenue, driving revenue and overall business growth.

I am also certified by LinkedIn as a digital marketing expert.
I love great ideas and uniqueness, I love driving results with tenacity and focus. I love Figures☺. With my solid marketing knowledge, I have helped several companies scale-up their sales and online presence with the use of digital marketing tools and analytics. I have a passion to drive change and create impacts in the lives of people.

Portfolio Outline

Introduction

As a Data Analyst working for a Charity Education for All. I have been 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 the database
- Increase the donation frequency of the donors.
- Increase the value of donations in the database.

In two weeks my team is having a fundraising strategy meeting for the following year, and I need to present insights from the donation data to inform your fundraising strategy and increase donations. I'm expected to Use the datasets EFO_Donation_Data and EFO_Donor_Data to answer the business problem.

Using Root-Cause Analysis, Descriptive and prescriptive analyses, I was able to highlight some of the KPIs to provide insights and solutions to this task.

Root Cause Analysis

Root-cause Analysis is a kind of analysis that seeks to reveal the real cause of a particular problem. The objective is to address the problem from its root so that the same problem do not keep-on reoccurring.

The Process

The first stage of carrying out a Root-cause analysis is understanding the problem. From this task, the problem of the Charity organization is that the value of Donation in the Database is not up to what the desired and the charity wants more donations.

The 5–WHY Approach

The 5-WHY approach helps us to dig deep into the root-cause of a problem by asking **'the why questions**' 5 times. The questions and process of carrying out the analyses and visualizations are outlined below:

Given the two dataset, since we have a primary and foreign key **(ID)**, I joined the two dataset using the inner-join command.

Select * FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id

After joining both tables, there were a total of 1000 rows and 16 columns in the new table.

To solve this business problem, it is important to examine the new table and develop critical business questions that can provide insight into the problem the organization wants to tackle.

First, I highlighted the important columns that are relevant to this analysis:

Important columns to be used

- First name
- Last name
- Gender
- Job field
- Donation
- state
- University
- Car

THE " WHY" QUESTIONS & ANSWERS

1. Why is the Donation value is not up to what we want in the organization ?

Answer: Donation value is not evenly distributed among the 49 States. Some states are donating far higher than others.

2. Why is the Donation value is not evenly distributed among the 49 States ?

Answer: Approximately 40% of the entire donations are coming from 5 states

Top 5 states by Donation

California	30,264
Texas	24,097
Florida	20,562
New York	14,759
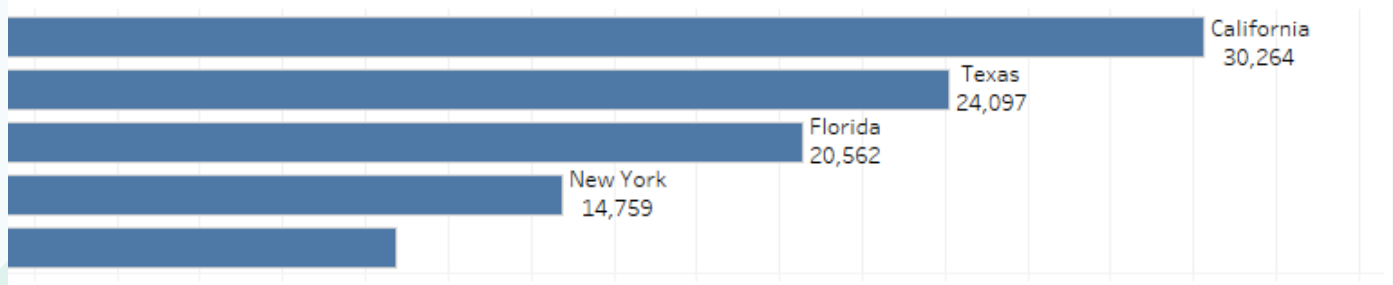
state Grouped	% of Total Donation along state Grouped	Distinct count of State	Donation
Others	59.68%	44	148,653
Top 5	40.32%	5	100,432

3. Why is 40% of the entire donations coming from just 5 states out of 49 states ?

Answer: This could either be because of the Population in those states or the average donation per person in those states is very high.

By examining the Average donation per person in the top 5 states and others we have:

Top 5 states = $254.7 per donor
Other states = $245.3 per donor

state Group..	Number of States	Avg. Donation per state
Others	44.0	245.7
Top 5	5.0	254.3

Average: The average donation per person among the top 5 states is about $254.3 while other states is about $245.7. There's a difference of about $9 which is not so large. However, I also examined the population difference before making a conclusion as to the major cause.

Population: - The population of the Top 5 states also took approximately 40% of the entire population of donors.

state Grouped	Number of States	Number of Donors	% of Donors
Others	44.0	605.0	60.50%
Top 5	5.0	395.0	39.50%

From these findings, we can see that the Top 5 states took about 40% of the entire donors, leaving 44 states to share the remaining 60%.

4. Why is the population in the top 5 states taking about 40% of the entire population ?

Answer: To find answer to this, let's examine the average number of donors in **EACH** state **contained** in the 2 groups **(Top 5 states and others).**

state Grouped	Number of States	avg number of donors per state
Others	44.00	13.75
Top 5	5.00	79.00

From this analysis we can see that the average number of donors in each state for the top 5 states is 79 donors **per state** while other states have **13 donors per state.**

At this point, we can establish that we have gotten to the root cause of why the value of donation in the database is not as much as what is desired.

The average only gives a central tendency, from the dataset, it could even be observed that states like wyoming, Maine, South Dekota have just 1 donor, some have 3 donors etc.

A strong solution to this challenge will be to create programs aimed at increasing the number of donors in other states asides the top states.

Other Insights & SQL Codes

In my quest to dig deeper into other analysis, I came up with important business questions for each section. They are outlined below:

A. Questions (Aim is to increase the value of donations in the database)

1. Which states are the most donations coming from (top 3)
2. What are the first 5 occupations that contributed the most to the Total donation
3. What are the occupations of the top 10 donors ?
4. What are the top 10 average donation per university

B. Questions (Aim is to increase the frequency of the donors)

1. What is the donation frequency of the top 10 donors and what kind of jobs do they do ?
2. What is the average donation amount per frequency ?
3. What are the frequencies of the least 10 donors and what kind of jobs do they do ?

C. Questions (Aim is to Increase the number of donors in the database)

1. Which gender is more likely to donate ?
2. Which donation frequency is convenient for most people?
3. Which state houses the most number of donors

SQL QUERIES & VIZUALIZATIONS

A. Questions (Aim is to increase the value of donations in the database)
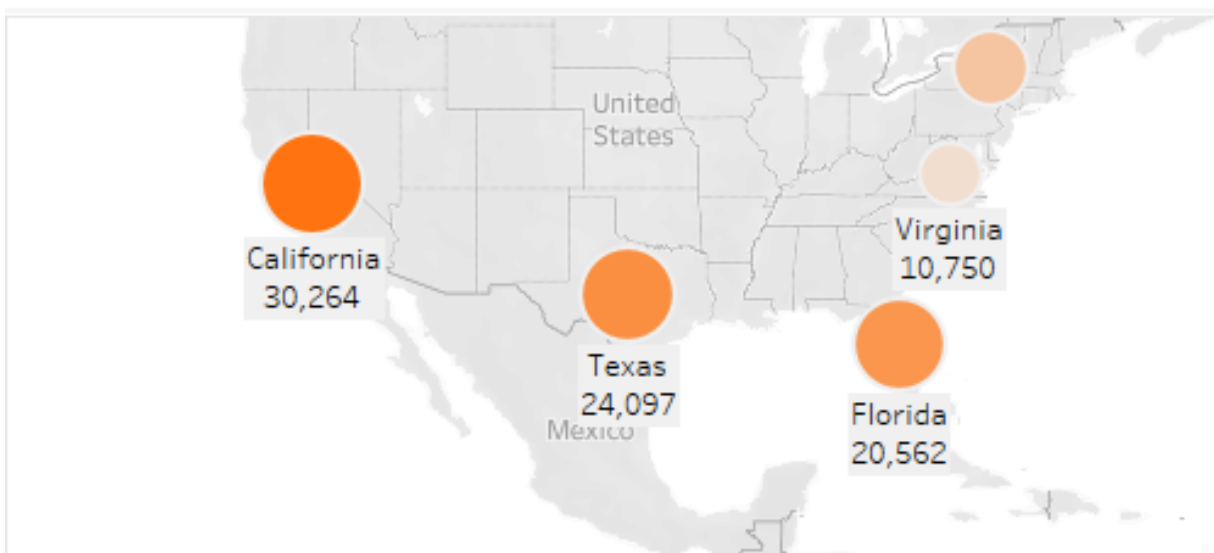
1. Which states are the most donations coming from (top 3)

Select state, SUM(donation) as state_donation
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
Group BY state
ORDER BY state_donation DESC
LIMIT 3

**California 30264**
Texas   24097
Florida   20562

California, Texas and Florida have the highest donations. This is as a result of the population In those regions because these are also the top 3 by count.

2. **What are the first 5 occupations that contributed the most to the Total donation**

Select job_field, SUM(donation) as job_donation
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
Group BY job_field
ORDER BY job_donation DESC
LIMIT 5

1. Human Resources
2. Research and Development
3. Product Management
4. Business Development
5. Engineering
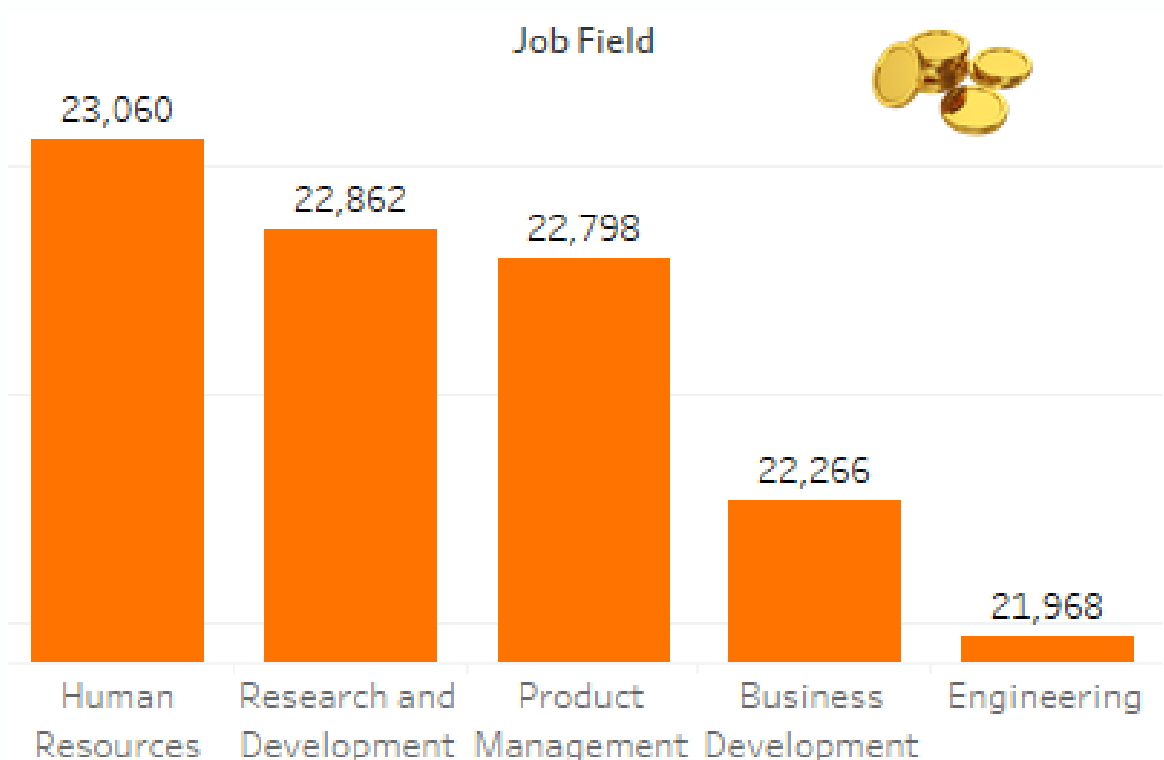
3. What are the occupations of the top 10 donors ?

The top 10 donors had occupations that spread across different fields. If there were clusters among the first 10. We might have assumed that certain kinds of jobs attract more donations but this doesn't seem to be the case..

Select FIRST_name, job_field, donation
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
ORDER BY donation DESC
LIMIT 10

Top 10 Donors by jobs

sundaytolu4@gmail.com

4. What are the top 10 average donations per university ?

This will enable us to know the universities that are contributing much to the total donations.

Select university, AVG(donation) as university_avg
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
GROUP BY university
ORDER BY university_avg DESC
limit 10

The top 10 universities have their average donations between 493 - 500 per university with walasik and leithgoe topping the list.
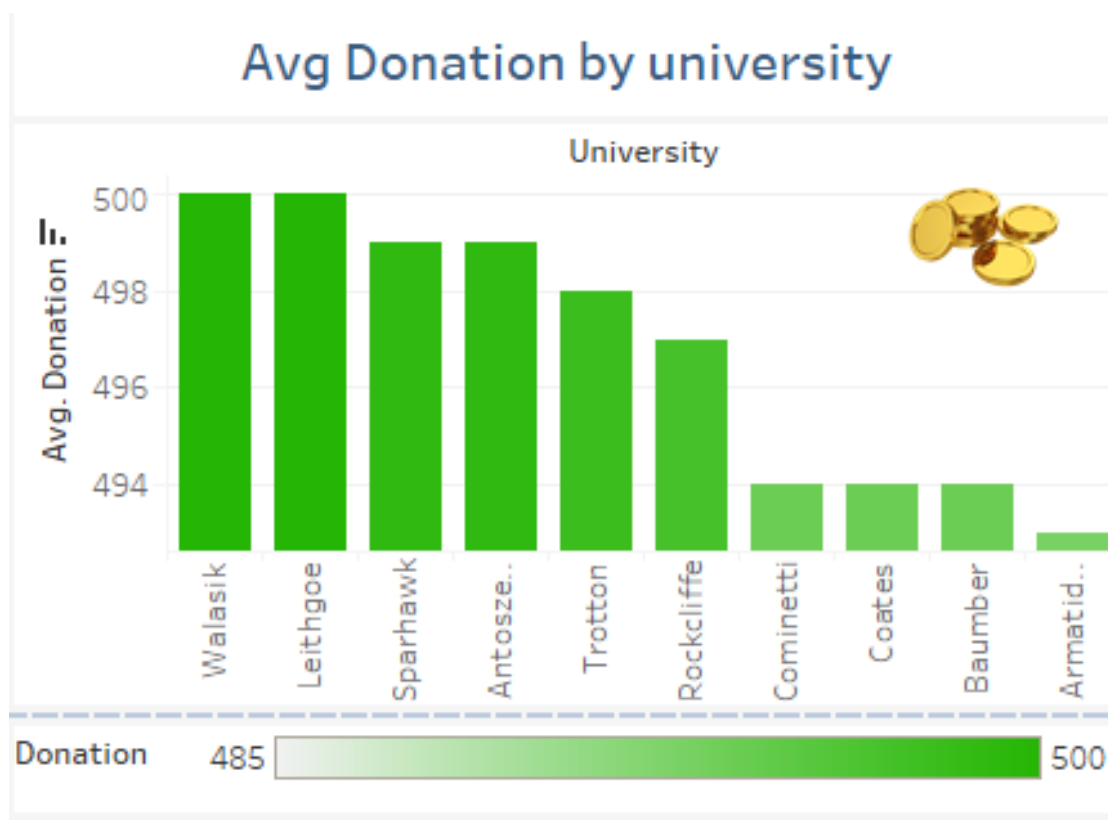

Avg Donation by university

sundaytolu4@gmail.com

B. Questions (Aim is to increase the frequency of the donors)

1. **What is the donation frequency of the top 10 donors and what kind of jobs do they do ?**

Select first_name, job_field, donation_frequency, donation
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
ORDER BY donation DESC
limit 10

50% of the top 10 donors pay monthly, 30% pay yearly and 20% pay weekly however, there are about 6 job fields among the top 10.. So it is safe to say that the top donors are spread across several jobfieds but most of them pay monthly

Top 10 Donors by Donation frequency
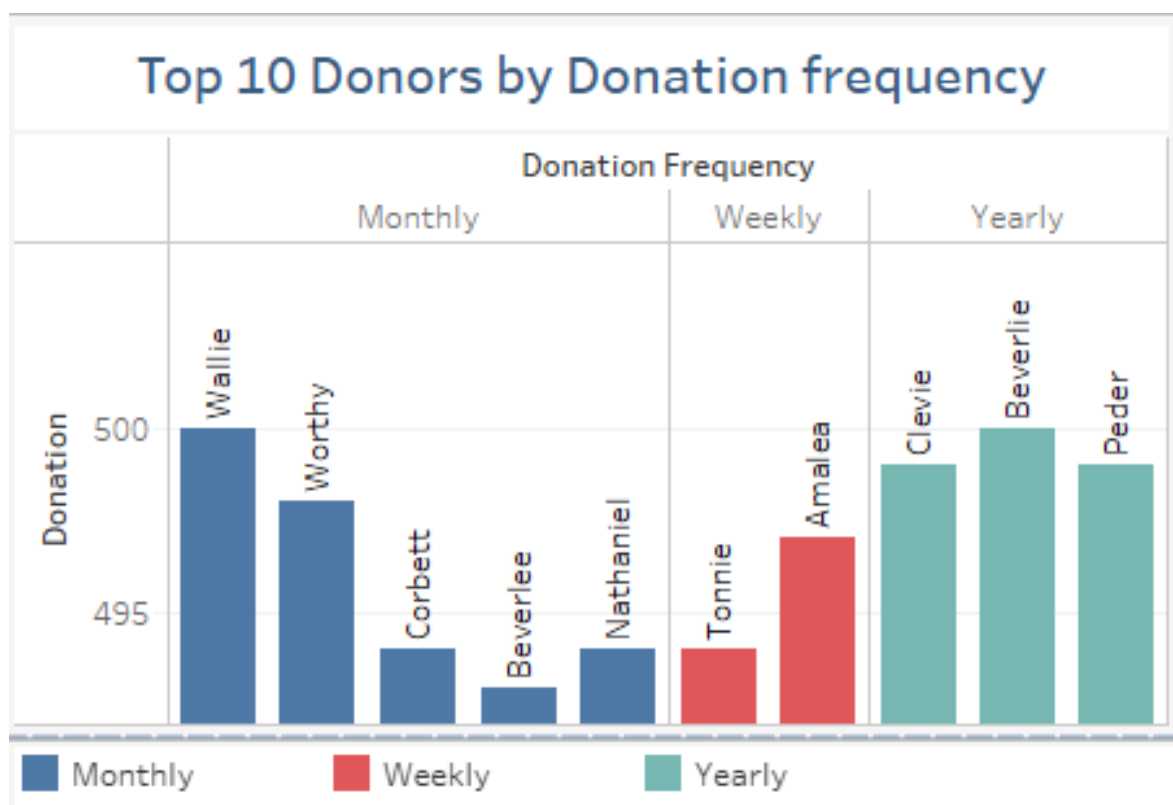
2. What are the frequencies of the least 20 donors and what kind of jobs do they do ?

Select first_name, donation_frequency, job_field, donation
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
ORDER BY donation
limit 20
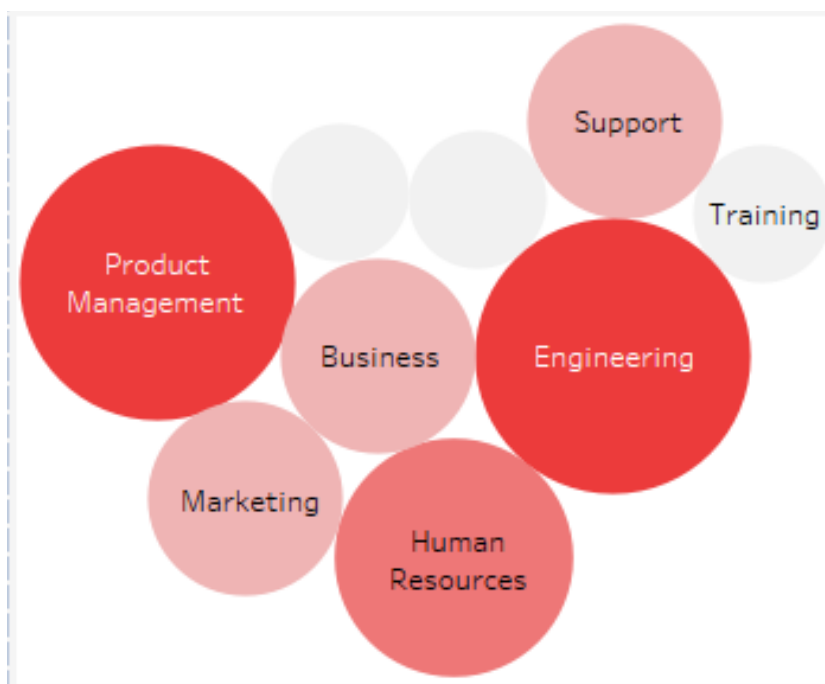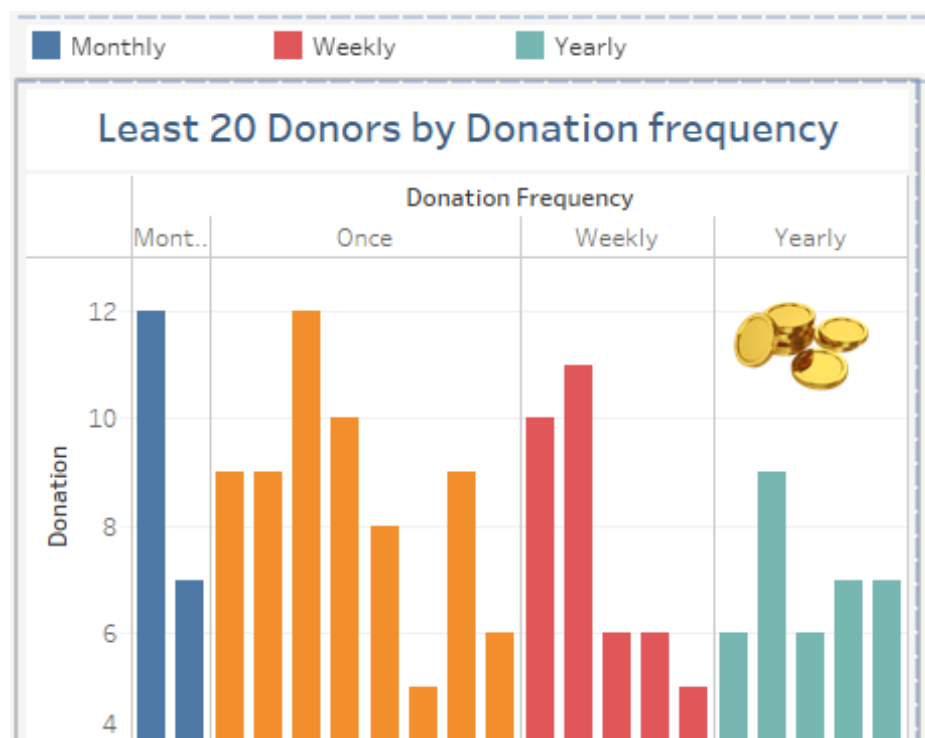
40% of the least 20 donors pay once, the remaining 60% is shared across weekly, monthly and yearly donors. However, their jobs are spread across several fields. This re-establishes the finding in A3 that the job fields are not really a determinant of the value of the donation made.
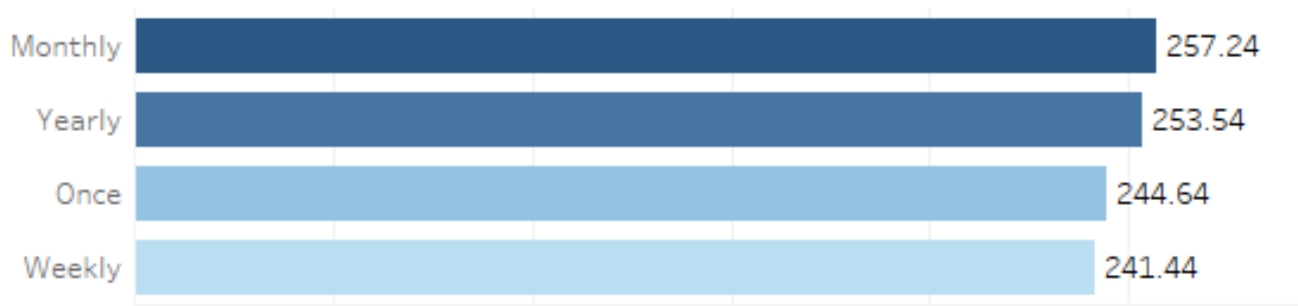
3. What is the average donation amount per frequency ?

Select donation_frequency, AVG(donation)
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
GROUP BY donation_frequency
ORDER BY donation DESC

Monthly —-- 257.24
Yearly —--- 253.54
Once —--- 244.64
Weekly —--- 241.4

Monthly	257.24
Yearly	253.54
Once	244.64
Weekly	241.44

The average donation per frequency shows that more money is donated per person on monthly or yearly basis than weekly.

C. Questions (Aim is to Increase the number of donors in the database)

1. Which gender is more likely to donate ?

We can find this by knowing the number of male and female donors in comparison with their average donation per gender.

COUNT OF GENDER

Select gender, COUNT(gender)
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
GROUP BY gender

Females: 508
Males : 492

AVERAGE DONATION

Select gender, AVG(donation)
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
GROUP BY gender

Female: 239
Male : 259

Despite the fact that the number of females were more than the males, the average donation per male was still higher than females. This means males have more propensity to donate than females and efforts should be made to reach out to more males to donate.

Gender by count

Gender by Average Donation

2. Which donation frequency is convenient for most people ?

We can know the most convenient donation frequency for people by knowing which donation frequency people prefered to use the most and compare with the average per frequency.

sundaytolu4@gmail.com

Select donation_frequency, COUNT(donation_frequency),
AVG(donation)
FROM Donation_Data
join Donor_Data2
on Donation_Data.id = Donor_Data2.id
GROUP BY donation_frequency
ORDER BY COUNT(donation_frequency) DESC
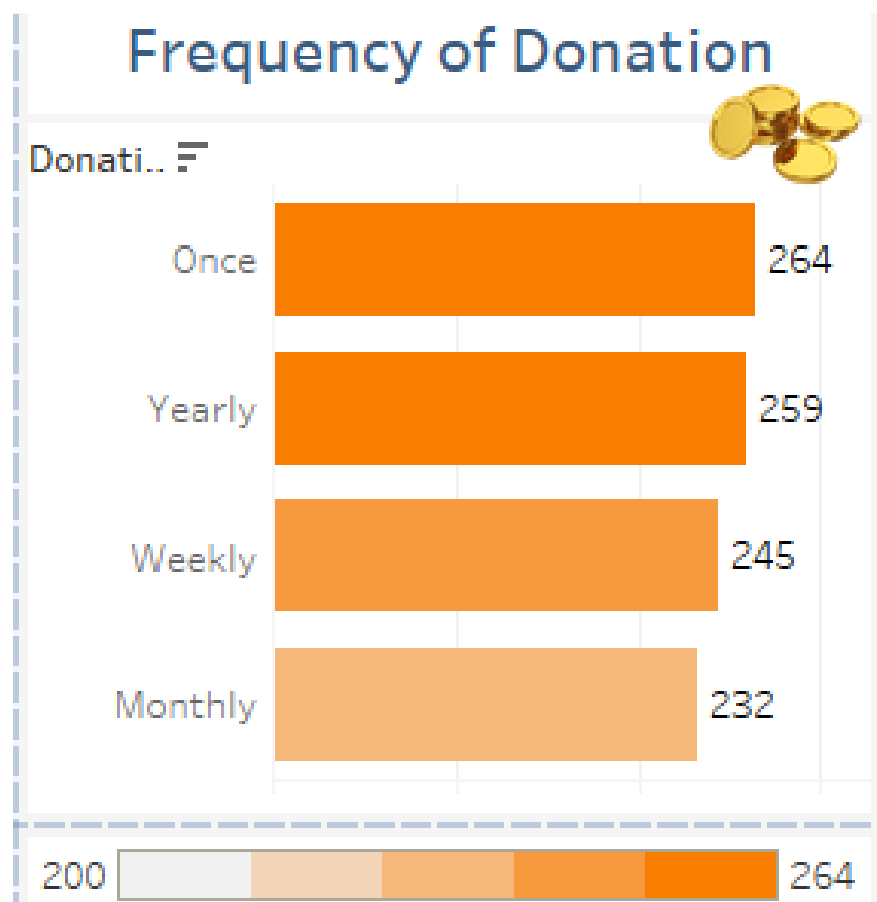
Frequency of Donation

Donati.. ⊟

Frequency	Count
Once	264
Yearly	259
Weekly	245
Monthly	232

200   264

From the results, most people choose to donate once and
this looks convenient for most people.

25%   50%   75%

Recommendations From Findings

1. Since the company is looking into increasing the value of donations in the database, it will be advisable to put more efforts into getting donors from other states that do not have so much records.

2. The average donation per frequency shows that more money is donated per person on monthly or yearly basis than weekly. To increase the value of donation in the database, the organization should encourage people to give their donation monthly or yearly.

3. From the results, most people choose to donate once and this looks convenient for most people. However, it may not be advisable to encourage people to donate once because people donating monthly on the average donated more in terms of value added to the contribution., so efforts should be put in place to encourage monthly donations

4. California has the highest number of donors followed by Texas and Florida. If the company is looking into increasing the number of donors in the database, The company could also put more efforts into developing strategies to scale-up the top 3 states since they have more population to explore.

Conclusion

The project examined some details of the donors in the database with the aim of increasing the value and number of donors.

The Root-cause analysis examined the root cause of the challenge and we found out that the main cause is the uneven distribution of the population in some states.

However, more quantitative analysis could have been made if other details of donors were provided like income range per annum.

Be that as it may, Some key recommendations have been made and with this, The charity should be able to increase the value of the donation in the database