

PROFESSIONAL BACKGROUND

My name is Chiko Liu Abdulsamad and I am a data analyst. I graduated from Ahmadu Bello University, Zaria, Nigeria in 2013 and have a bachelor's degree in Mechanical Engineering. I started my career as an instructor in computer science in 2014 and in the following year, I went on to sales and marketing in a telecommunication company where I worked for the next 2 years. I stumbled into data analysis when I stepped in to solve a database management problem with no expertise on ground to solve it. In 2017, I then went on to become a public health analyst in the expanded program on immunization in the north-eastern part of Nigeria for 4 years, during this time, I created a database for my designated territory to track all the immunization data, created a capacity building template for health worker to track their baseline knowledge on the immunization program as well as an assessment to track their progress. I later went on in 2021 as an independent consultant to design a result management system using google sheet instead of building an application for a tech-based school in other to save cost. I am a tech Enthusiast, a problem solver, a quick learner and I required minimum supervision to work. I just concluded my internship in data science using python and I am currently studying online to be a data analyst using SQL and Tableau. I am proficient in the using of python, Excel, SQL, Kobo Collect tool and a novice with Tableau.

Contents

PROFESSIONAL BACKGROUND	1
INTRODUCTION	3
ROOT CAUSE ANALYSIS PROCESS	4
Problem:.....	4
Questions:	4
INSIGHTS FROM ANALYSIS.....	5
Here are some of the insights gotten from the analysis of the dataset	5
1. Analysis by gender:	5
2. Analysis by state.....	6
3. Analysis by Movie_genre	7
4. Analysis by Car types.....	8
5. Analysis by Job fields.....	9
6. Dashboard.....	10
FINDINGS AND RECOMMENDATIONS.....	11

INTRODUCTION

A database of donors and their donations was given for analysis with the objectives of increasing number of donors, their frequency of donations as well as the amount they donated.

Analysis was done based on gender, types of jobs, the state, type of cars driven and shirt size. The reason for these choices is based on physical observation without necessary having contact with the donors. Others were dropped due to incomplete dataset such as university and personal choices such as types of movies they watched.

At the end of the analysis:

1. based on gender only, male donors gave more compared to their female counterparts.
2. California, Texas, Florida, New York and Virginia gave more donations compared to the rest of the states in the USA
3. Although job fields were also a consideration, but there were no significant differences between the maximum and minimum amount donated based on the job field however there was a significant difference between the preferred movie types of donors and their donations as donors who preferred drama, comedy, documentary, drama|romance donated far more than those who preferred other types of movies.
4. Donors with large shirt size gave significantly more than those with extra small shirt size.

ROOT CAUSE ANALYSIS PROCESS

Problem:

To Increase regular donations through fundraising.

Questions:

Why is there a database to track donors and their donations?

Why seek to increase the donations?

Why fund education?

Why are there people who don't have access to education?

Why are there people who do not have sponsors for their education?

INSIGHTS FROM ANALYSIS

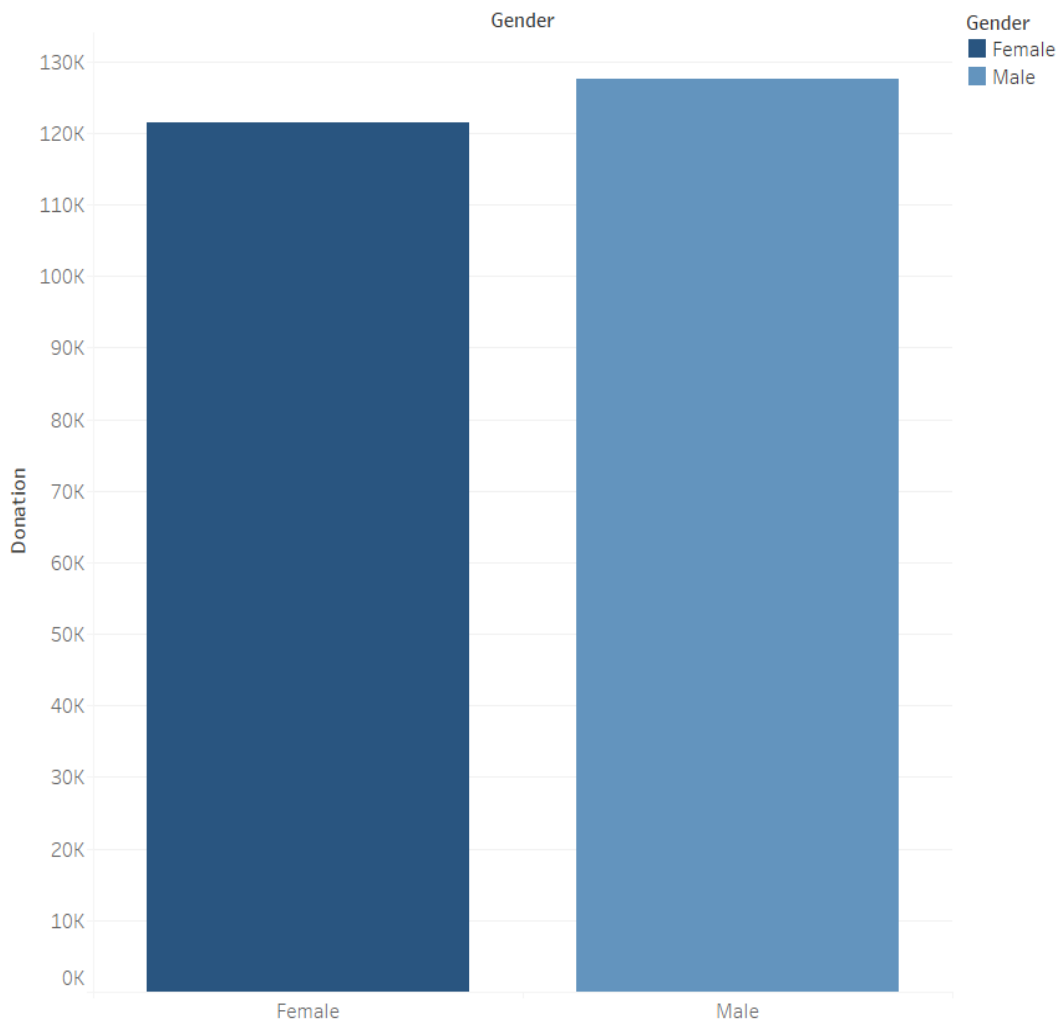
Here are some of the insights gotten from the analysis of the dataset.

1. Analysis by gender:

Generally male donors gave more donations than their female counterparts

```
SELECT gender, count(donation) AS "FREQ", SUM (donation) as "Total", ROUND  
(AVG (donation),0) as "Average"  
FROM portfolio_csv pc  
GROUP BY gender  
ORDER BY FREQ DESC;
```

Gender



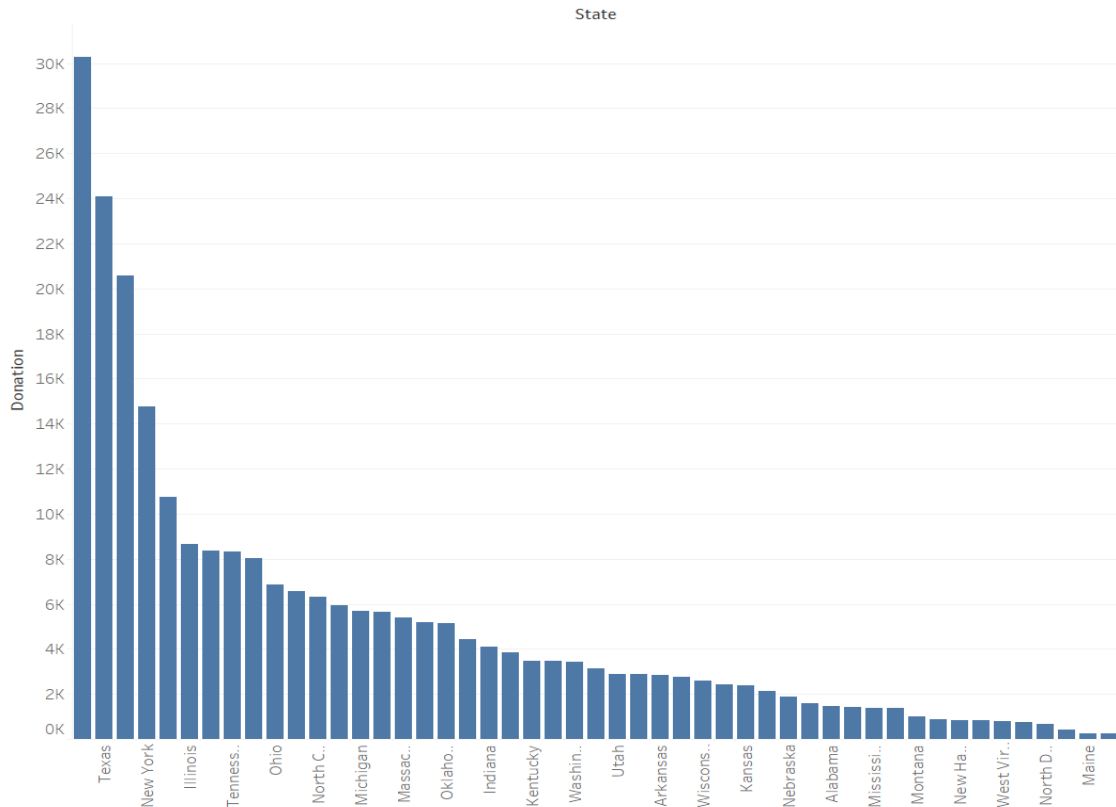
Sum of Donation for each Gender. Color shows details about Gender.

2. Analysis by state

States like California, Texas, Florida, New York and Virginia gave more donations than the rest of the states.

```
SELECT state, SUM (donation) as 'total donation', COUNT (donation) as '# of donors', round (AVG (donation),0) as 'average amount donated'
FROM portfolio_csv pc
group by state
ORDER BY SUM (donation) DESC;
```

Donation by state



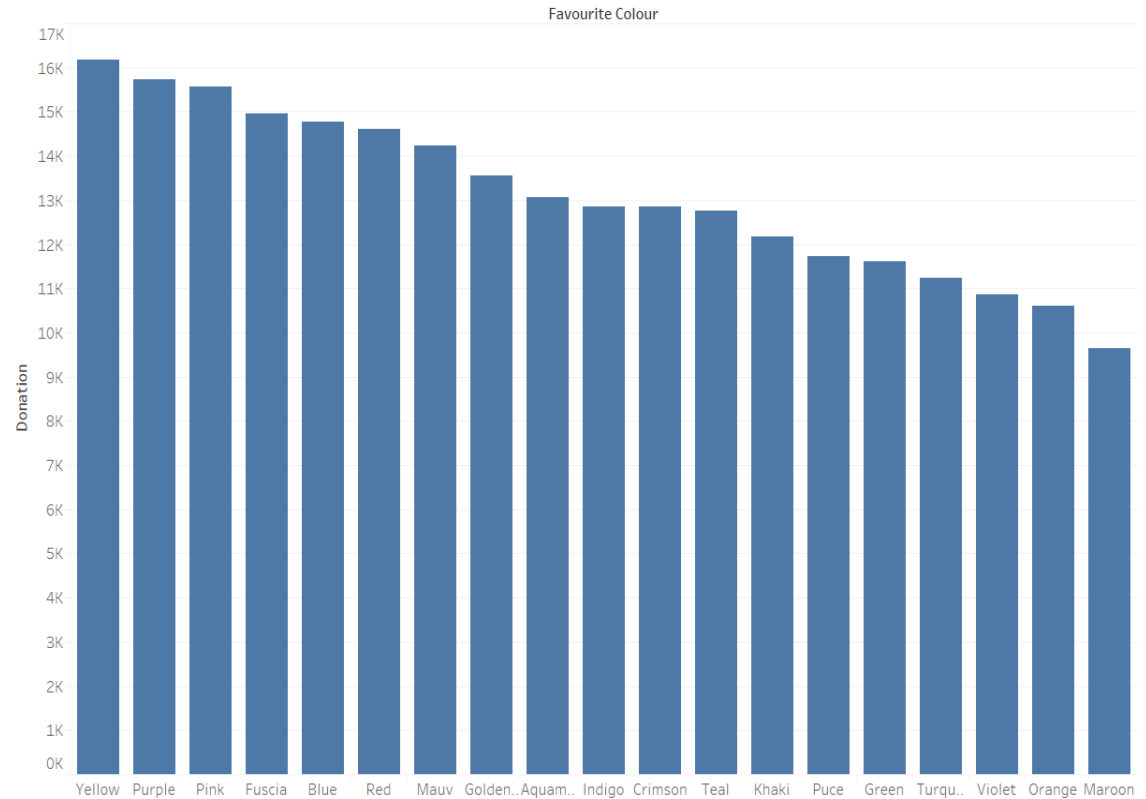
Sum of Donation for each State. The data is filtered on Donation Frequency, which keeps 8 of 8 members.

3. Analysis by Movie_genre

Donors who watch Drama, Comedy and Documentary gave more when compared to the rest of the movie genre

```
SELECT movie_genre, COUNT (donation_frequency) as "# of donors", SUM
(donation) as "Total Donation"
FROM portfolio_csv pc
GROUP BY movie_genre
ORDER by "Total Donation" DESC;
```

Color

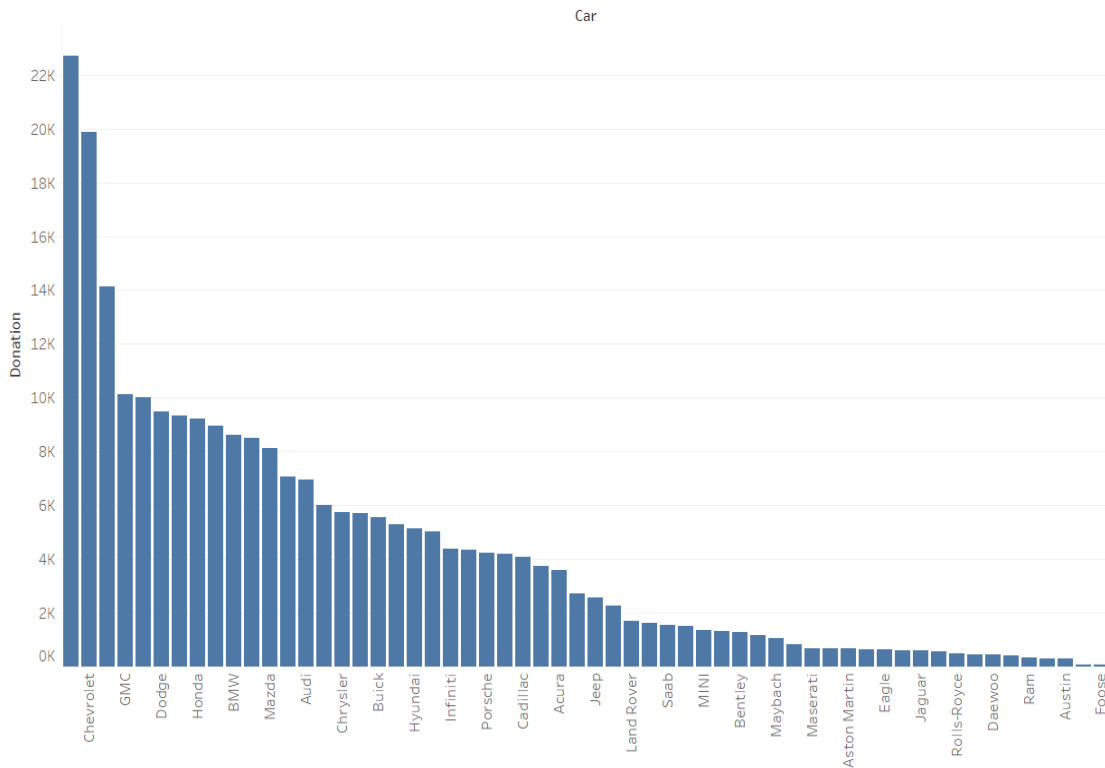


4. Analysis by Car types

Donors with expensive cars like Lamborghini, Porsche etc. weren't much on the donor's database and gave significantly less when compared to donor with cars like Ford, Toyota, Chevrolet etc. who were quite high in number and gave significantly more.

```
SELECT car, COUNT (car) as "# of cars", SUM (donation) as "Total Donation",  
ROUND (AVG (donation),0) as "Average Donation"  
from portfolio_csv pc  
group by car  
ORDER BY "Total Donation" DESC;
```

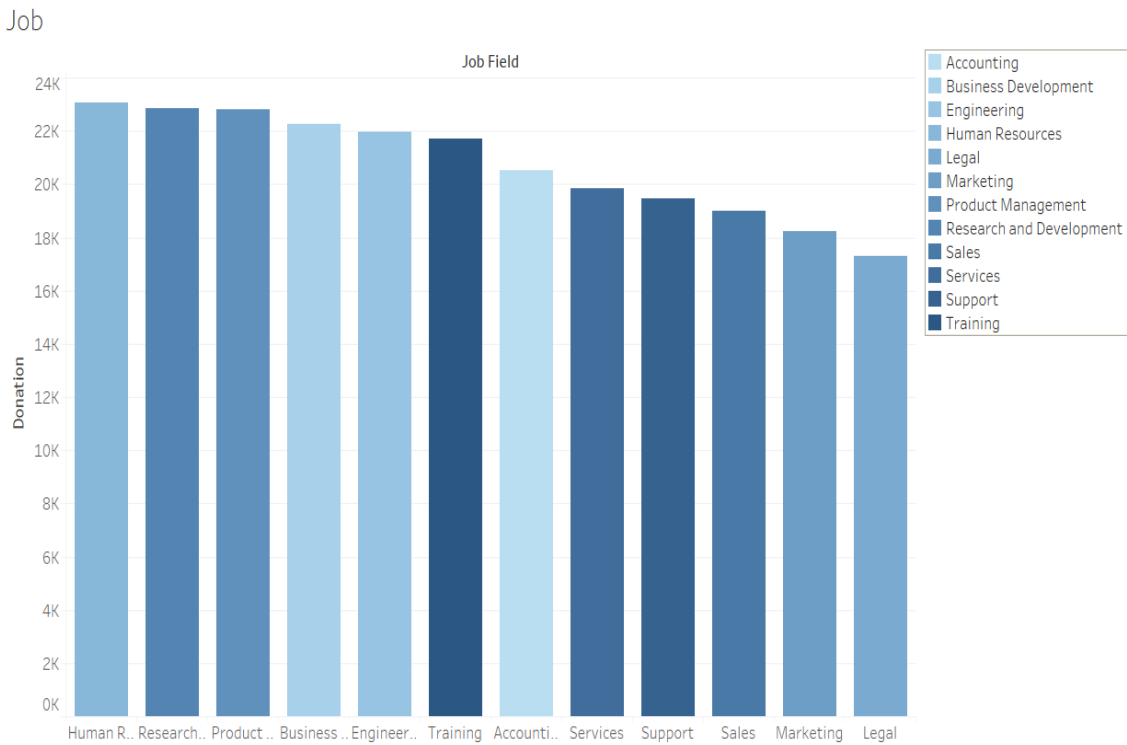
Cars



Sum of Donation for each Car.

5. Analysis by Job fields

```
SELECT job_field , COUNT(donation_frequency) as "# of Donors", SUM(donation)
as "Total Donation"
from portfolio_csv pc
GROUP BY job_field
ORDER BY "Total Donation" DESC;
```

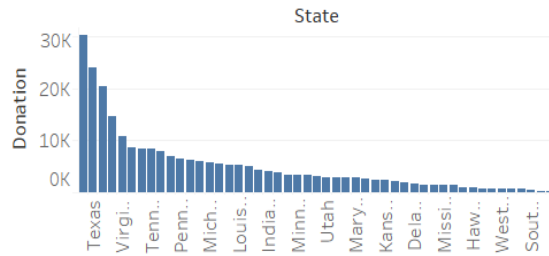


Sum of Donation for each Job Field. Color shows details about Job Field.

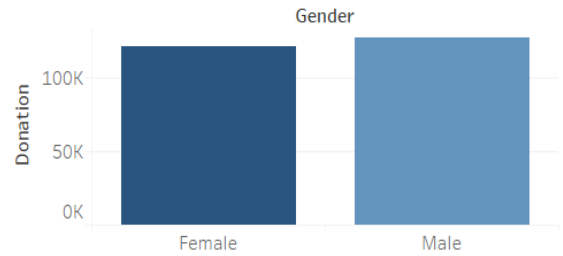
6. Dashboard

Below is a dashboard of the analysis done above

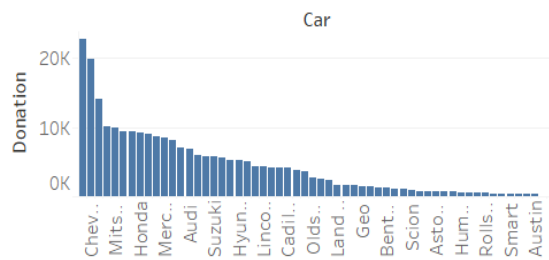
Donation by state



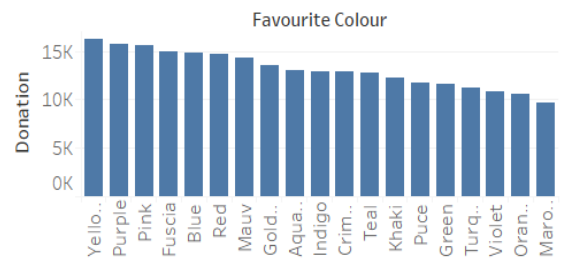
Gender



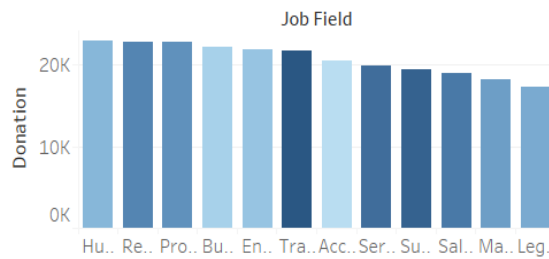
Cars



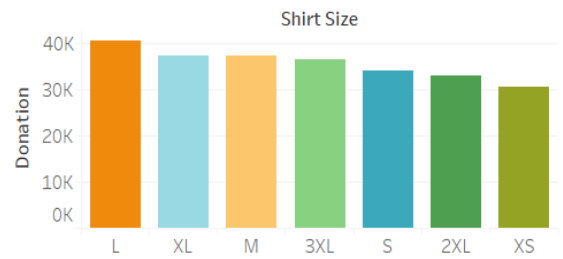
Color



Job



ShirtSize



FINDINGS AND RECOMMENDATIONS

1. Null values: the null values in university, second language can be as a result of infinite reasons and that in itself can result to infinite implications, therefore the data collection system has to be able to account for Null values either by capturing most of the data or discard the data if the null values are much.
2. Redundant datasets such as Car type wasn't helpful in bringing useful insights as donor who had expensive cars such as Mercedes-Benz, Lamborghini, Rolls-Royce donated little compared donors who had Ford, Toyotas and Chevrolet.
3. The dataset didn't give any information about the timeframe the dataset was collected for i.e., dataset after 1year or 2021 donations or donation dataset between 2019 – 2021 thus making the donation frequency ineffective thus anytime an analysis is be done on any dataset, a timeframe of the dataset should be captured.
4. Missing datasets e.g., Marital Status, Age etc. these datasets if available might have given more insights as to the reason why men gave more than women, donors with certain high earning jobs donated more or less than expected based on their expected salary, why donors with luxurious cars gave significantly less than those with other brands of vehicle.
5. The total donations in the state of California, Texas, New York, Virginia were higher when compared to Hawaii, Alaska, North Dakota, Wyoming and this begs the question if the latter states had awareness of the organization and its objectives, therefore there is the need to focus awareness and sensitization of **Education for All** in the states that donated significantly less.