

Introduction

For the task at hand, we're given data for states in Nigeria. The data was initially presented as three different files and spread over a course of three years, 2018-2020. The data is for the state and federal earnings in three measures: FAAC (federal allocation), VAT (a component of FAAC) and IGR (state revenue).

Before we begin, two important things to point out about the dataset:

- The data for the IGR stops at the halfway mark of 2020. The reasoning for this is unknown. However, extrapolating the data or filling into it ourselves, would not be feasible or recommended. Inserting a volume as huge as half a year's worth would cause more issues and inaccuracies than benefits. Therefore, we'll assume it to be the full data for the year.
- Also, the trends are likely to decrease for the year 2020 owing to the COVID-19 pandemic. We will move through the analysis keeping this in mind and therefore, unless there is a strong intuition for a negative trend in say, VAT or IGR, the reason would most likely be COVID-19.

For the sake of this analysis, we've decided to dive into the way each state has been performing on a federal and individual level. Overall, the idea is to find the worst and the top performing states. This data and insight can be useful for a variety of stakeholders.

For the federal governments, this could help adjust allocation. Basically, if a state is more independent i.e. the IGR is high and growing, the resources allocated to them could be turned towards states that need more of them. In the same way, the vice versa would also work.

This same insight can prove invaluable for business owners, especially those starting out. If a state has a high IGR income, it may correlate to them having a higher % of fees and charges that fall into the purview. Given that some of that may be relevant to certain businesses, this insight can help them properly decide a base of operations.

The use-cases are endless, and these are speculative ideas, of course.

That said, before any analysis, the key task was to merge the data into a sort of master dataset which would ideally look like the following.

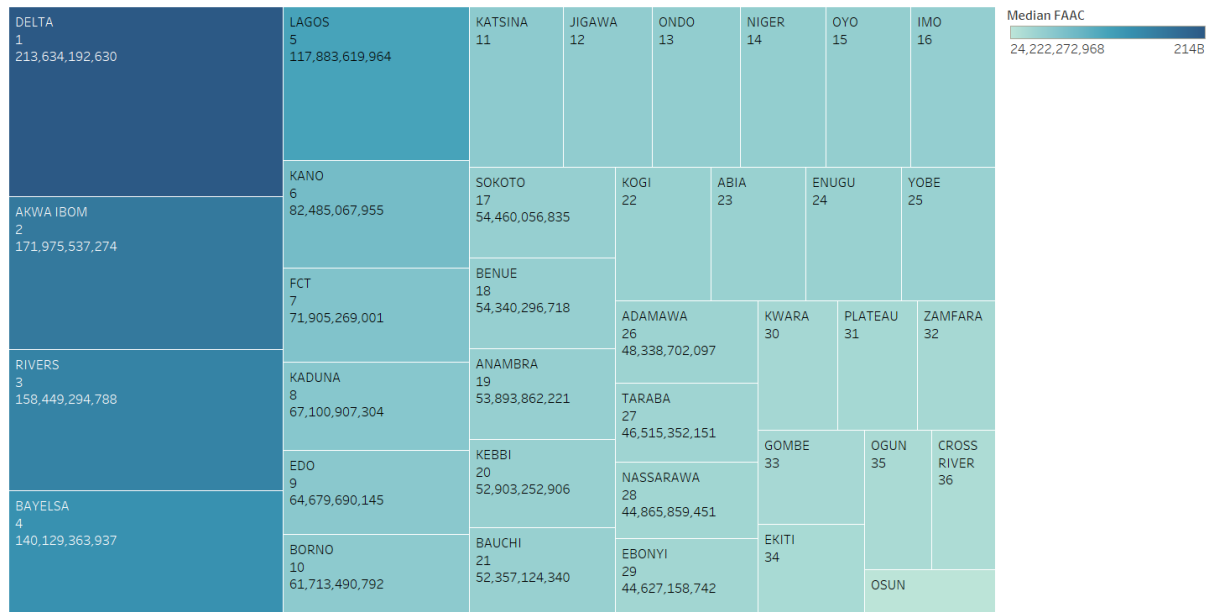
| STATE | TYPE | YEAR | REVENUE |
|-------|------|------|---------|
|-------|------|------|---------|

This would make it easy to analyse and group the data based on any of the available factors. This task and other aggregation were performed in RStudio. The aggregated data was then exported as .csv files and visualized using Tableau.

We'll walk through the five analyses one-by-one.

1. Median FAAC

Median FAAC (2018-2020)



State, Rank of Median Faac and sum of Median Faac. Colour shows sum of Median Faac. Size shows sum of Median Faac. The marks are labelled by State, Rank of Median Faac and sum of Median Faac.

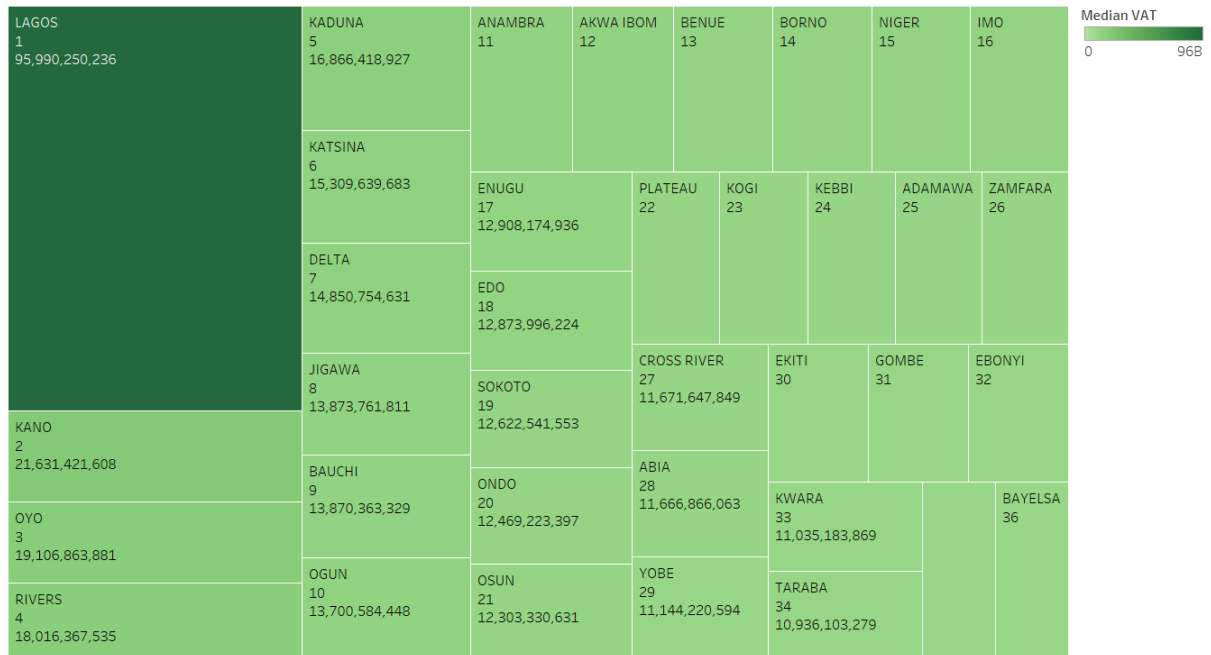
For the first analysis, we're trying to understand which are the top-5 states in terms of their median Federal Allocation over the given time-frame. This would mean that these states are highly dependent on the Federal government.

It's clear from our Tree Map that the highest of the five states are Delta, Akwa Ibom, Rivers, Bayelsa, and Lagos. Similarly, the states with the lowest median FAAC are Gomba, Ekiti, Ogun, Cross River, Osun.

Note: The state's high FAAC over three years does not mean that they are solely dependent on the government, however. The idea is to understand if FAAC can be better allocated. We'll try to understand this with further visualisations.

2. Median VAT

Median VAT (2018-2020)



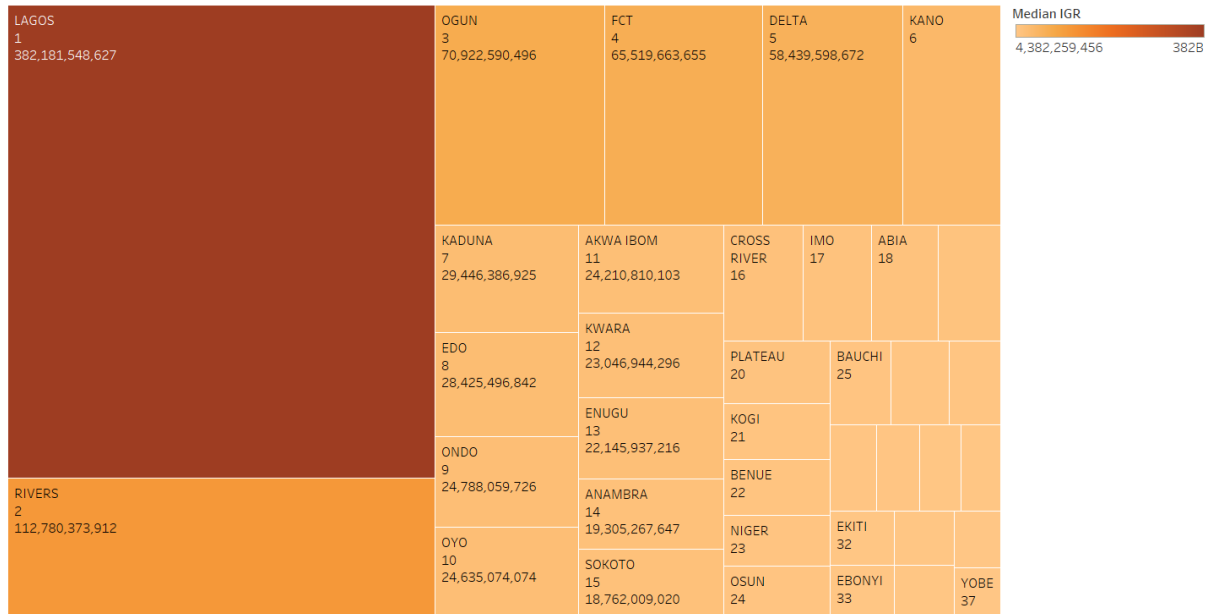
State, Rank of Median Vat and sum of Median Vat. Colour shows sum of Median Vat. Size shows sum of Median Vat. The marks are labelled by State, Rank of Median Vat and sum of Median Vat.

As discussed earlier, VAT is a component of the FAAC. Our rankings for median VAT put things a bit in perspective when viewed directly with the FAAC visualization. If we notice, on a median, Lagos had the highest FAAC component.

This suggests that Lagos has provided more business since VAT is dependent on the businesses performing in the state. (Logical leap to be verified based on a domain expert's opinion). Other components may or may not be high for Lagos.

3. Median IGR

Median IGR (2018-2020)



State, Rank of Median Igr and sum of Median Igr. Colour shows sum of Median Igr. Size shows sum of Median Igr. The marks are labelled by State, Rank of Median Igr and sum of Median Igr.

From this visualization, there's a kind of obvious idea rising that one, Lagos is probably a business hub. From a quick lookup, it also happens to be the most densely populated area. Therefore, one can conclude with the high IGR that while Lagos gains much from in-state business and VAT components, the federal government can allocate some of the other components to other states.

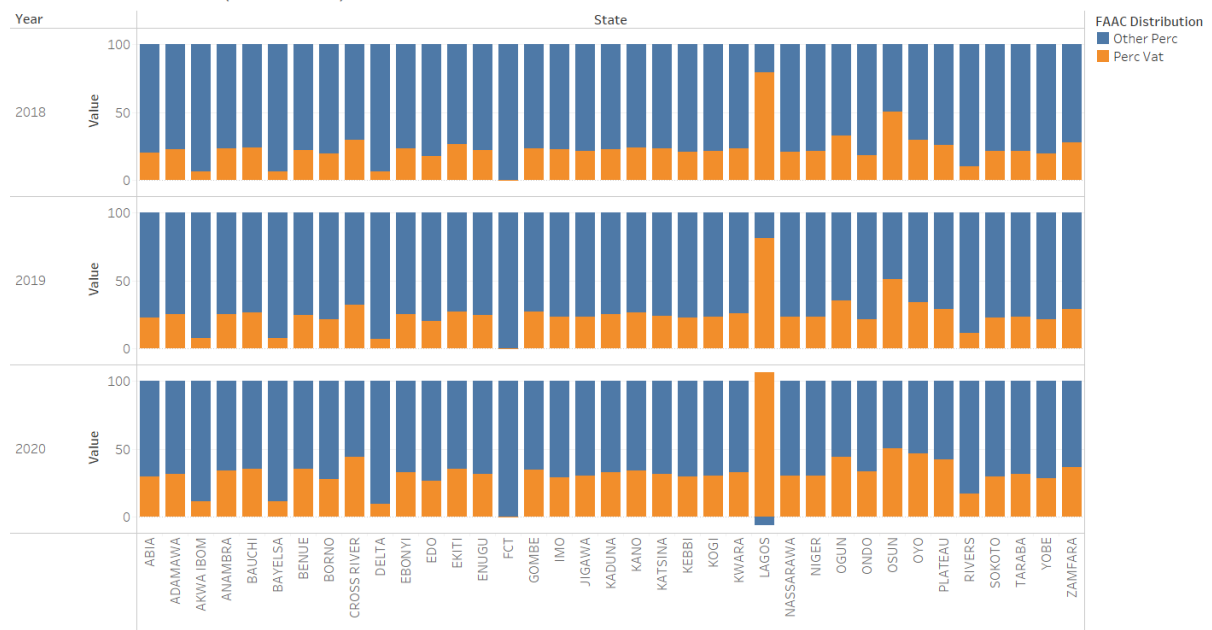
Delta, Rivers and others also fall into similar categories. Ideally, states that are well to do by themselves should receive less support in Other Components. This could lead to a holistic development of the entire country in time.

Also, for business holding stakeholders, it's clear that places like Lagos, Delta and Rivers are the ideal hotspots to start businesses since high VAT and IGR medians indicate that these states have great businesses.

For the next visualization, we'll explore changes in the VAT v Other components of the FAAC.

4. VAT v Other % Change

FAAC Distribution % (2018-2020)



Other Perc and Perc Vat for each State broken down by Year. Colour shows details about Other Perc and Perc Vat.

Here's a distribution of the % change for VAT and Other components over the course of three years. Basically, we're trying to understand whether a state is VAT-dominant. A quick-glance shows a few things.

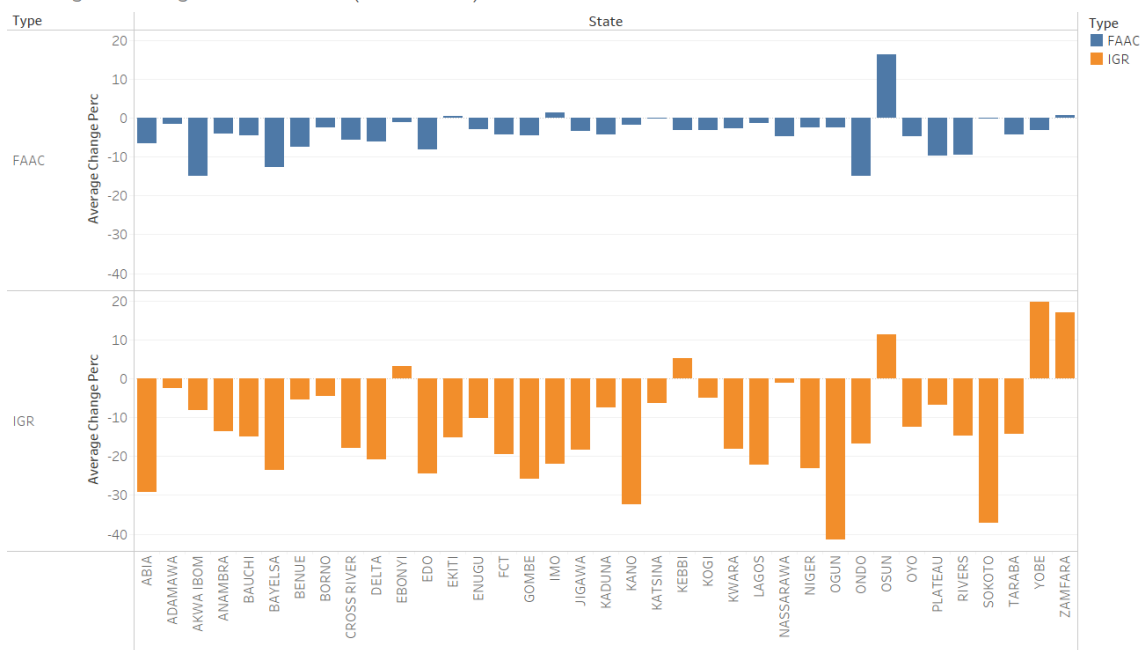
The foremost thing to notice is that by 2020, the % of VAT share in the FAAC has increased. This could be due to a number of reasons such as policies, business booming or VAT rates.

This also shows that places like FCT, Akwa Ibom, Bayelsa, Delta, and FCT are almost absolutely dependent on the Other (non-VAT) allocations of the FAAC.

Please note, the discrepancy of the bar in Lagos is because there's a -6.3% component for Other. This could be due an error in the dataset or perhaps, some recording/analysis error that was not rectified. For merit, let us ignore the Lagos' bar for 2020 until we have better data.

5. Mean % Change FAAC & IGR

Average % Change in FAAC & IGR (2018-2020)



Sum of Average Change Perc for each State broken down by Type. Colour shows details about Type. The view is filtered on Type, which keeps FAAC and IGR.

Here, we're trying to analyse and compare the average % change (2018-2019, 2019-2020) for both the FAAC and the IGR. The idea is to understand a state's individual performance over the two years for the % change in their IGR and FAAC.

A high negative average % change in IGR over the course of two years can indicate either a worst performing year (probably owing to COVID-19) or a consistently lowering business acumen.

In any case, this analysis can be drilled down further to extract more insight as to how each state might have performed in an ideal scenario and which states exhibit a consistently negative % change for the IGR.

On the other hand, a high negative % change in FAAC indicates that the policies and the federal government are now favouring the specific state less which could either be owed to shifting priorities or a state becoming self-sufficient.

Of course, the data and context warrants a deeper analysis.