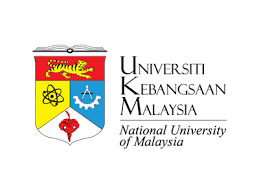
**AWS Project report**

Federal Government Finance

Dr. Ahmad Tarmizi Abdul Ghani

TTTU6314

BUSINESS INTELLIGENCE AND ANALYTICS



AFRAA MOHAMMED GAASHAN

P143085

2024-2025

1. Get any data set from Data.Gov.MY (make sure the data you choose has lookup value data provided.

The dataset is like the following

1-The dataset name

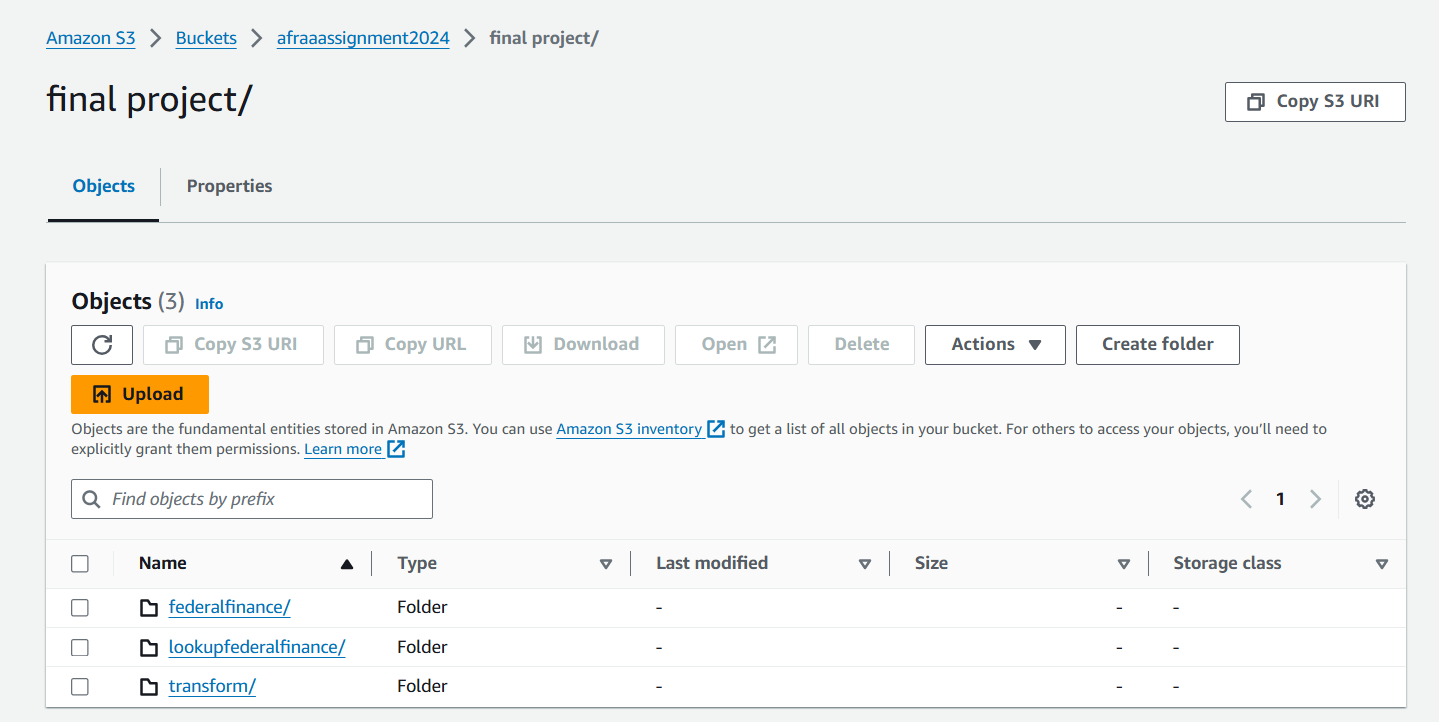
Quarterly Federal Government Finance; Quarterly revenue, operating and development expenditure, overall surplus/deficit, and sources of financing for the Federal Government of Malaysia.

2-Lookup Table: Federal Finance

If specified, the lookup table will be left-joined against the category and variable name in the other Federal Government Finance datasets.

2. Save the data and lookup value in S3.

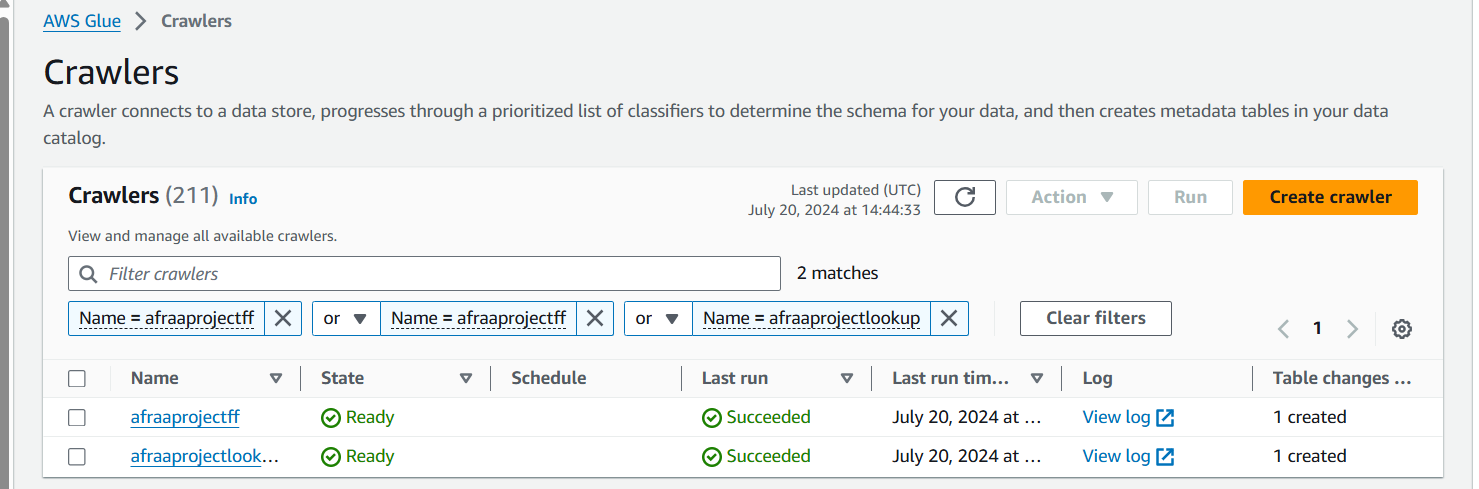
In bucket afraaassignment2024 in a folder named (final project), there are the folders containing CSV files



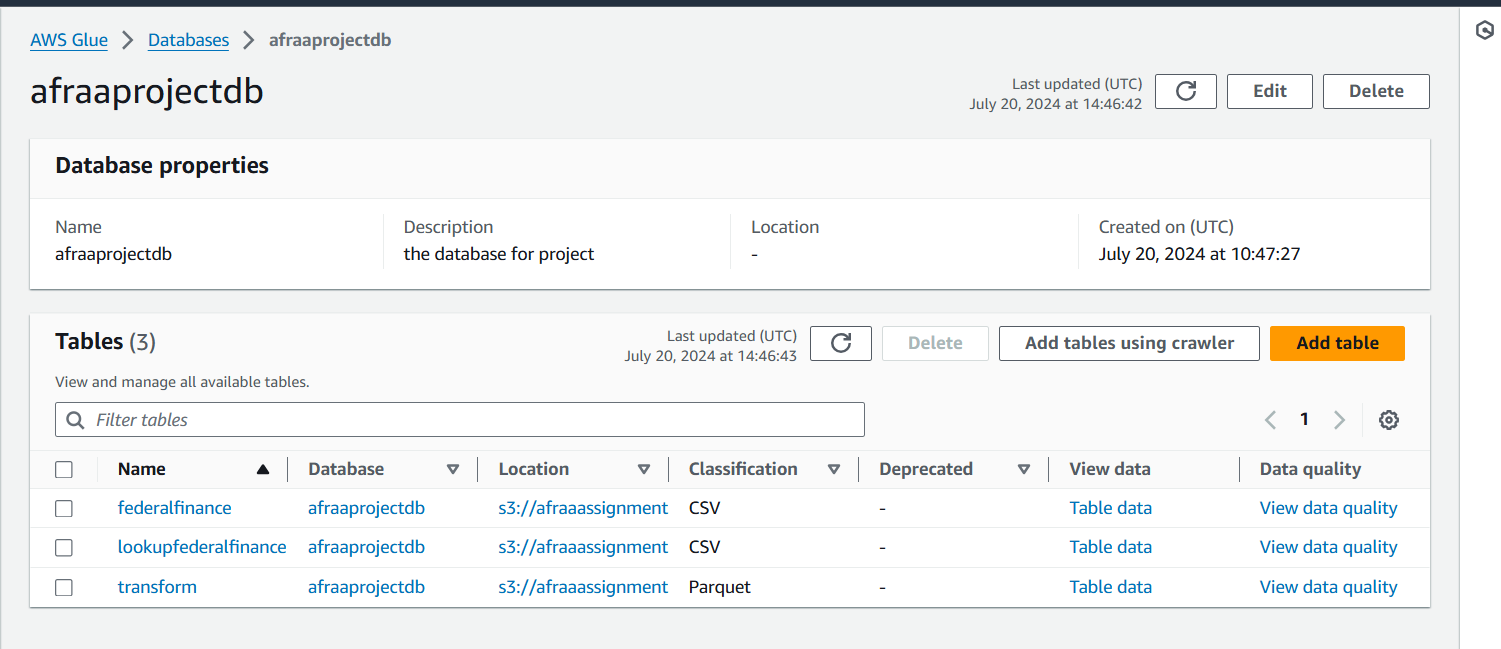
3. Do crawling to the data to create a data catalog using Glue.

Both two have to do the same steps for crawling

1. go to the data catalog then crawl
2. create a crawler by selecting the data source then the database etc…
3. then go to the database to see the datasets and tables



1. to make sure that the crawler runs successfully with data we have to see the database and make some query



The queries

A screenshot of a computer

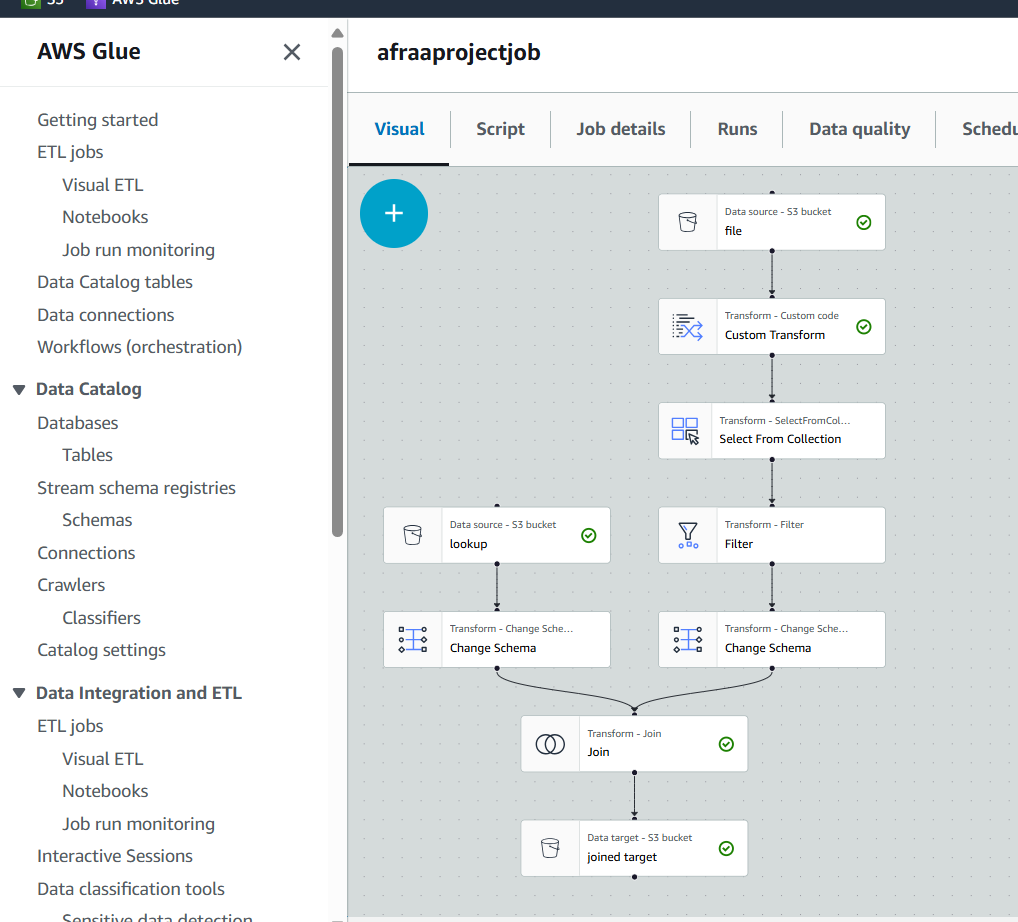
Description automatically generated

A screenshot of a computer

Description automatically generated

4. Do Glue ETL to the data to Join the data with the lookup value.

1- in ETL jobs create a new one and apply to join for transformation



2- click save then runs to run the job

A screenshot of a computer

Description automatically generated

3- make crawling for transform data and save it in a different folder

A screenshot of a computer

Description automatically generated

4- see the job result

A screenshot of a computer

Description automatically generated

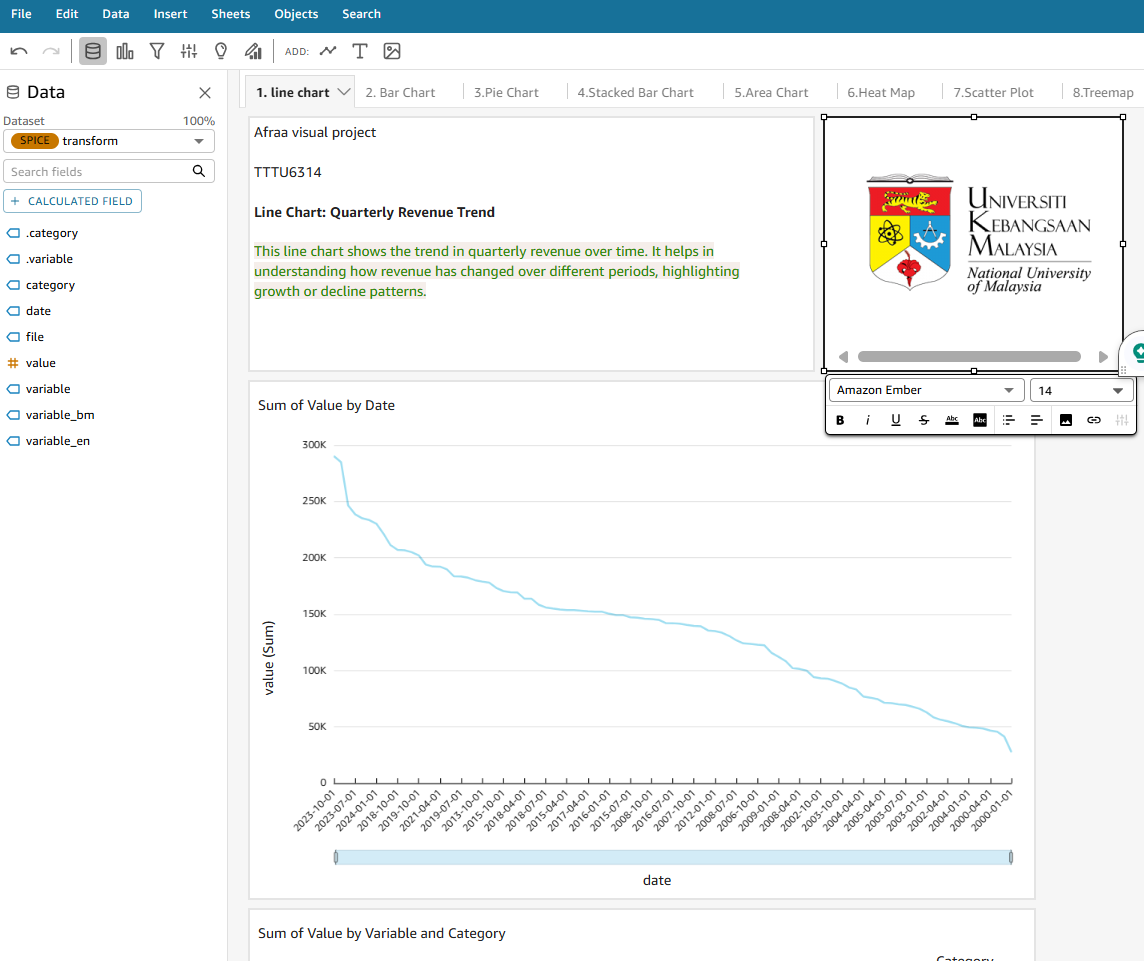
5- apply query

A screenshot of a computer

Description automatically generated

5. Visualize the JOINED data in Quicksight

1- line chart for Quarterly revenue trend

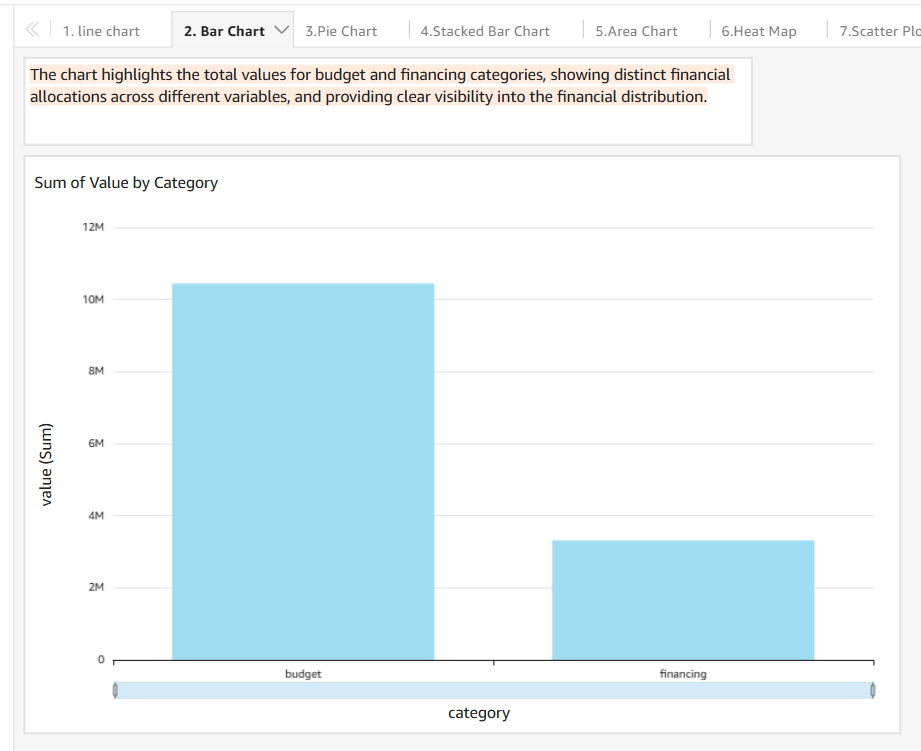


1. Sum of Value by Date; The total value consistently decreases over time, showing a significant decline in financial metrics from 1996 to 1999.

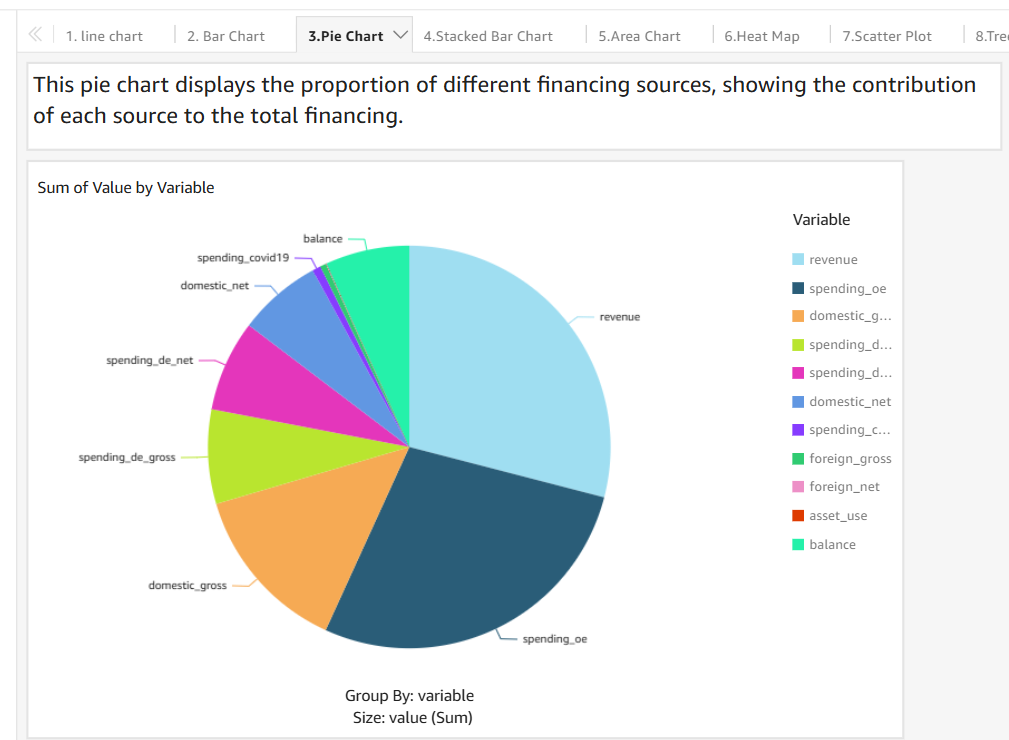
2. Sum of Value by Variable and Category; The budget category exhibits more variability compared to financing, indicating fluctuating budget allocations during this period.

2-bar chart for Category-wise expenditure

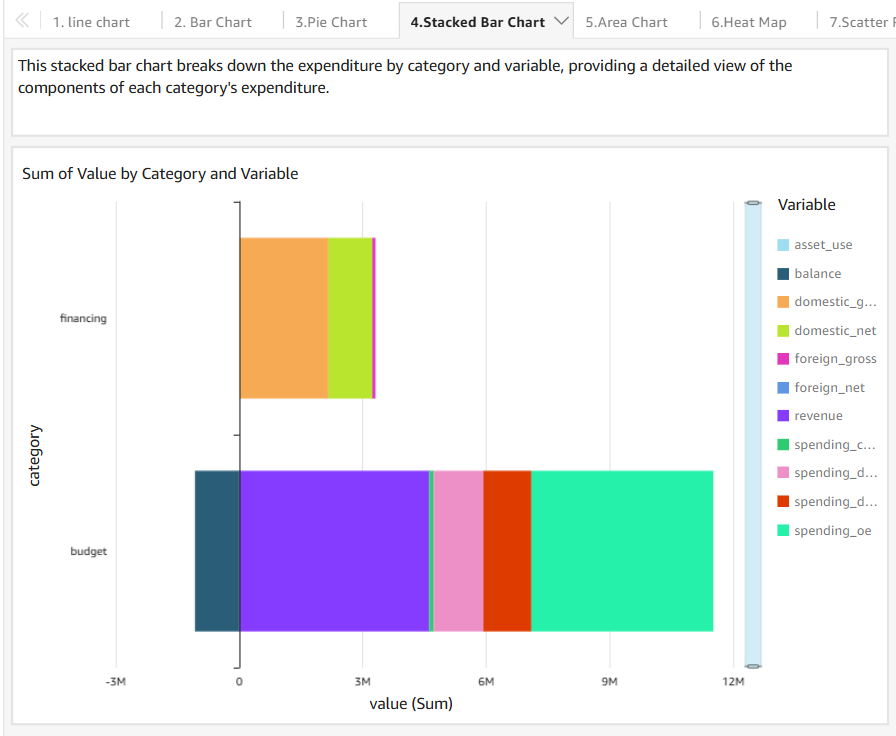
It clearly displays the sum of values for each category (budget and financing) while showing the breakdown by different variables. This bar chart illustrates the expenditure for each category, identifying the highest and lowest spending categories.



3- Pie Chart for Proportion of financing sources



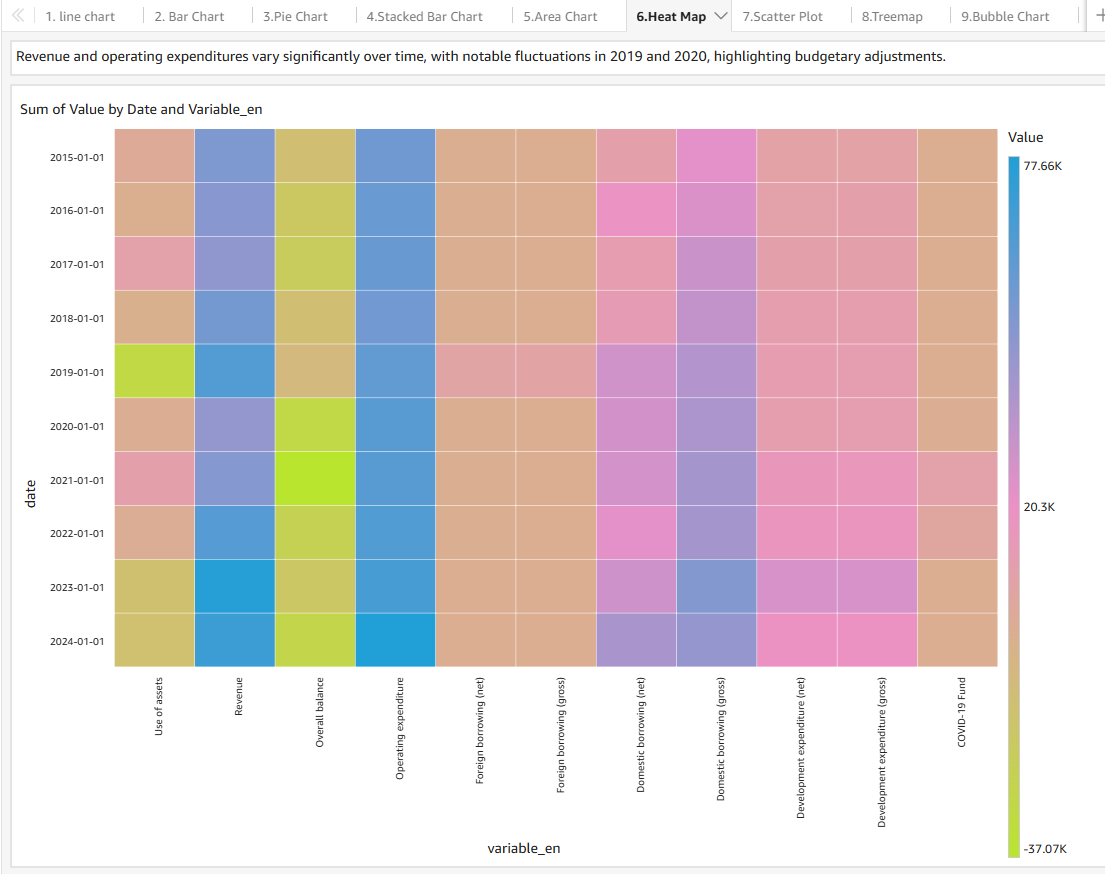
4-Stacked Bar Chart for Expenditure Breakdown by Category

 5- Area Chart for Surplus/Deficit Over Time

A screenshot of a graph

Description automatically generated

6- Heat Map for quarterly financial performance



The heat map shows the total sum of values across different variables and dates, with the color intensity representing the magnitude of these values. By examining the color patterns, it's evident that certain variables such as revenue and operating expenditures exhibit significant changes, particularly in the years 2019 and 2020. This suggests notable financial adjustments or events during these periods. The visual differentiation allows for a quick assessment of how various financial components have evolved over time, indicating key periods of budgetary shifts or economic activities.

7-scatter plot should show the relationship between revenue and operating expenditure. If done correctly, you should not need to manually enter values, as the calculated fields will automatically isolate the required data.

A screenshot of a graph

Description automatically generated

8- Treemap for Hierarchical View of Expenditure

A screenshot of a computer

Description automatically generated

9- Bubble Chart to impact of different variables on financial Performance

A Sankey diagram is particularly useful for understanding the distribution and flow of values between different nodes.

A screenshot of a computer

Description automatically generated

- Waterfall Chart: Quarterly Changes in Financial Position10

A screen shot of a graph

Description automatically generated