HACKATHON 2 - GDP Sector Dominance

Tracker

Name : K Guru Charan

REG NO : RA2211026010141

DEPT : CSE AI&ML

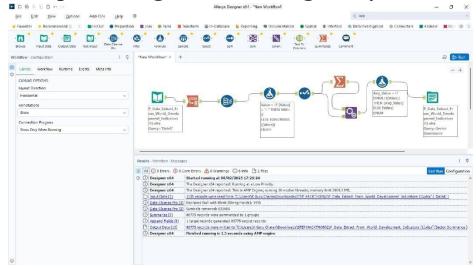
Introduction

TThis report covers the second core implementation module of the GDP360 project: the GDP Sector Dominance Tracker. I have had the honor of being a Data Analyst with the World Economic Insights Council (WEIC), and this module is a vital part of how we reflect on the transformation of global economies that have taken place over the past 50 years (1975–2025).

The report objectives are to accurately track and visualize the shift in the sectors of contribution to a country's Gross Domestic Product (GDP) by decade. As we note the dominant sector per decade for the major global economies, we gain critical insight into the structural changes to each economy and can see periods of major growth, decline, and transformation to sectors.

The report has a strong ETL pipeline for data structure and preparation built with Alteryx; SQL Common Table Expressions (CTEs) for transformations and analysis; and Power BI for interactive comparisons of histories and visualizations of sectors. The track and their comparisons will provide an informative snapshot of the major shifts in sectors, which have been and continue to be influenced. These results will hopefully help governments, economists, and think tanks understand how to quantify and relate practically and strategically how historical growth and development can impact economies in the future.

1. Preprocessing the data using alteryx



Alteryx Workflow: Sector Dominance Dataset Preparation

- Workflow Name: Sector Dominance Tracker.yxmd
- **Purpose**: To clean, transform, and impute missing values in sector-wise GDP contribution data for identifying dominant economic sectors across decades.

Data Sources:

- Connected to Excel file:
 - P_Data_Extract_From_World_Development_Indicators (1).xlsx
- Sheets/Queries Used:
 - Data\$ (Raw sectoral GDP data)
 - Sector Dominance (Processed output)

Tools Utilized

- Input Data Tool: To read raw sectoral GDP amounts across countries and years.
- Data Cleanse Pro Tool: To convert invalid blanks ("-") to nulls to cleanly handle them in datasets.
- Formula Tool: To convert text amounts to a numeric value; to handle nulls using logic;
- Summarize Tool: To produce average value to impute missing records.
- Append Fields Tool: To append average value to every row for imputation purpose.
- Formula Tool (2nd): To replace nulls in Value column with average values using corresponding sector/year combination.
- Output Data Tool: To write the final cleaned datasets to Excel.

Cleaning Logic:

- Changed placeholder symbols ("..") to proper nulls.
- Converted GDP contribution amounts that were text values to numeric values by utilizing TONUMBER().
- Imputed nulls with average sectoral values by group.
- Performed a clean, consistent schema of output datasets for SQL/BI analytics.

Final Output Dataset:

- Sheet: Sector Dominance
- Dataset in structured table format with cleaned Country, Sector, Year, and GDP Contribution (%).

Connection:

- Final dataset feeds into Power BI dashboards to illustrate the sector dominance described above.
- This also serves as the springboard to identify any sectoral shifts, changes and ongoing trends captured across decades.

Run Status:

- Successfully executed using AMP Engine in 1.3 seconds
- No errors or warnings reported.

2.Schema

1. Table: Dim Country

- Purpose: Stores unique country codes and names for dimensional analysis.
- · Columns:

```
Country_Code (PK, TEXT)
Country Name (TEXT)
```

2. Table: Fact GDP Data

- Purpose: Stores core GDP and high-level inflation metrics per country per year.
- · Columns:

```
Fact_ID (PK, INTEGER AUTOINCREMENT)
Country_Code
```

```
    (FK → Dim_Country.Country_Code, TEXT) o
    Year (INTEGER)
    Total GDP USD (REAL) o Inflation Pct (REAL)
```

3. Table: Fact Sector GDP

- **Purpose:** Stores unpivoted detailed sector contributions (percentage and calculated USD value) for each country, year, and sector.
- · Columns:

```
Country_Code (PK, FK → Dim_Country.Country_Code, TEXT) o

Year (PK, INTEGER)

Sector_Name (PK, TEXT) - e.g., 'Agriculture', 'Industry', 'Services',
 'Manufacturing' o Sector_Percentage (REAL)

Sector_GDP_USD (REAL) o Decade (INTEGER)

Dominant_Sector_Flag (BOOLEAN) -

(Optional, can be derived in Power BI)
```

4. Table: Fact Macro Indicators

- **Purpose:** Stores detailed macroeconomic indicators beyond core GDP and inflation, for impact analysis.
- · Columns:

```
Country_Code (PK, FK → Dim_Country.Country_Code, TEXT)

Year (PK, INTEGER)

Series_Code (PK, TEXT)

Inflation_Consumer_Prices_Annual_Pct

(REAL)

GDP_Growth_Pct (REAL)

Is_US_Recession (INTEGER or BOOLEAN)

Employment_Agriculture (REAL)

Employment_Industry (REAL)

Employment_Services (REAL)

Unemployment ILO Estimate (REAL)
```

5. Table: Staging Macro Data

• **Purpose:** Intermediate staging table for raw macroeconomic data during ETL.

Columns:

```
ID (PK, INTEGER)

Country_Name (TEXT)

Country_Code (TEXT)

Year

(INTEGER)

Total_GDP_USD (REAL)

Inflation Pct (REAL)
```

6. Table: Staging Sectors Agri Ind

• **Purpose:** Intermediate staging table for raw agriculture and industry sector data during ETL.

· Columns:

```
ID (PK, INTEGER)

Country_Name (TEXT)

Country_Code (TEXT)

Year (INTEGER)

Agriculture_Pct (REAL)

Industry Pct (REAL)
```

7. Table: Staging Sectors Serv Man

- **Purpose:** Intermediate staging table for raw services and manufacturing sector data during ETL.
- · Columns:

```
o ID (PK, INTEGER)

Country_Name (TEXT)

Country_Code (TEXT)

Year (INTEGER)

Services_Pct (REAL)

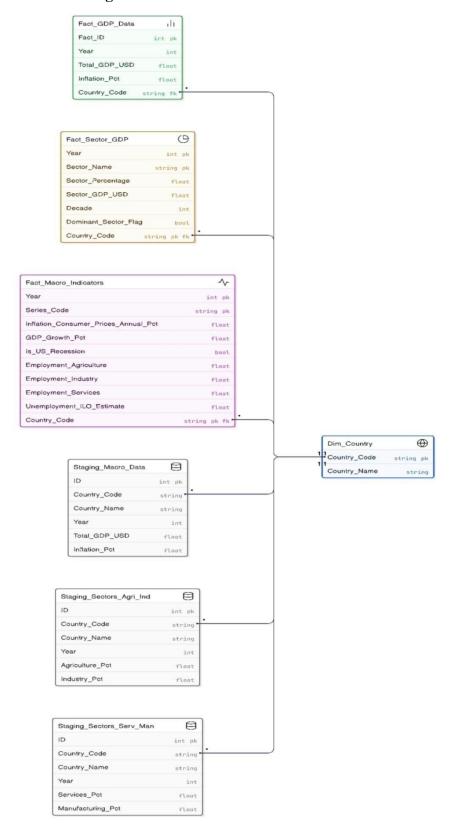
Manufacturing Pct (REAL)
```

Relationships:

• Fact GDP Data.Country Code → Dim Country.Country Code

- Fact_Sector_GDP.Country_Code → Dim_Country.Country_Code
- Fact_Macro_Indicators.Country_Code \rightarrow Dim_Country_Code

2.b Schema Diagram



3.SQL Queries

→ Joining and inserting the data

```
INSERT INTO Fact_GDP_Data (
    "Country_Code",
    "Year",
    "Total_GDP_USD",
    "Inflation_Pct",
    "Agriculture_Pct",
    "Industry_Pct",
    "Services_Pct",
    "Manufacturing_Pct"
  SELECT
    GDP3."Country_Code", -- The Country Code comes from the GDP3 table
    GDP3."Year",
    GDP3."gdp %",
    GDP3."Inflation, consumer prices (annual %)",
    GDP."Agriculture",
    GDP."Industry",
    GDP2."Services",
    GDP2."Manufacturing"
  FROM
    GDP3
  -- Join using Country Name and Year
  JOIN
    GDP ON GDP3."Country Name" = GDP."Country Name" AND GDP3."Year" = GDP."Year"
  JOIN
    GDP2 ON GDP3."Country Name" = GDP2."Country Name" AND GDP3."Year" = GDP2."Year";
```

Output:

	Total_GDP	Inflation_Pct	Agriculture_Pct	

→ Unpivoting the table

WITH Unpivoted_Sector_Data AS (

SELECT

Fact_ID,

Country_Code,

Year,

Total_GDP_USD,

'Agriculture' AS Sector_Name,

Agriculture_Pct AS Sector_Percentage

FROM Fact_GDP_Data

WHERE Agriculture_Pct IS NOT NULL -- Exclude rows where this specific sector's data is missing

UNION ALL

SELECT

Fact_ID,

Country_Code,

Year,

Total_GDP_USD,

'Industry' AS Sector_Name,

Industry_Pct AS Sector_Percentage

FROM Fact_GDP_Data

WHERE Industry_Pct IS NOT NULL

UNION ALL

OUTPUT:

Fact_ID	Country_Code	Year	Total_GDP_USD	Sector_Name	Sector_Percentage
602	AUS	1990	311427000000	Agriculture	4.193445564
910	AUS	1990	3146230000000	Agriculture	4.193445564
602	AUS	1990	311427000000	Industry	28.77344621
910	AUS	1990	3146230000000	Industry	28.77344621
602	AUS	1990	311427000000	Manufacturing	5.340542705
910	AUS	1990	3146230000000	Manufacturing	5.340542705
602	AUS	1990	311427000000	Services	59.54564211
910	AUS	1990	3146230000000	Services	59.54564211
587	AUS	1991	325975000000	Agriculture	3.171197981
911	AUS	1991	3046250000000	Agriculture	3.171197981
587	AUS	1991	325975000000	Industry	27.647039
911	AUS	1991	3046250000000	Industry	27.647039
587	AUS	1991	325975000000	Manufacturing	7.521622993
911	AUS	1991	3046250000000	Manufacturing	7.521622993

→ calculating the sector dominance

Avg_Sector_Percentage_In_Decade

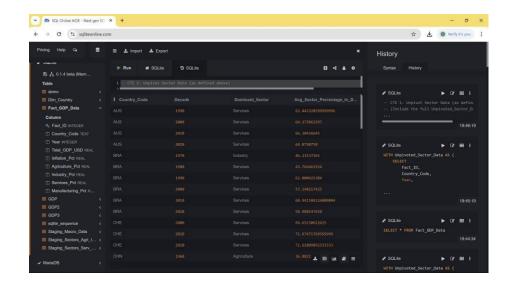
```
Unpivot Sector Data (as defined above)
WITH Unpivoted_Sector_Data AS (
  SELECT
    Fact ID, Country Code, Year, Total GDP USD,
    'Agriculture' AS Sector_Name, Agriculture_Pct AS Sector_Percentage FROM Fact_GDP_Data WHERE Agriculture_Pct IS NOT
NULL
  UNION ALL SELECT
    Fact_ID, Country_Code, Year, Total_GDP_USD,
    'Industry' AS Sector Name, Industry Pct AS Sector Percentage FROM Fact GDP Data WHERE Industry Pct IS NOT NULL
UNION ALL SELECT
    Fact_ID, Country_Code, Year, Total_GDP_USD,
    'Services' AS Sector Name, Services Pct AS Sector Percentage FROM Fact GDP Data WHERE Services Pct IS NOT NULL
UNION ALL SELECT
    Fact_ID, Country_Code, Year, Total_GDP_USD,
    'Manufacturing' AS Sector_Name, Manufacturing_Pct AS Sector_Percentage FROM Fact_GDP_Data WHERE Manufacturing_Pct IS NOT
NULL
),
Calculate Average Sector Contribution per Country, Decade, and Sector
Average_Sector_Contribution_Per_Decade AS (
  SELECT
    Country_Code,
    (Year / 10) * 10 AS Decade, -- This calculates the start of the decade (e.g., 1975 becomes 1970)
Sector Name,
    AVG(Sector Percentage) AS Avg Sector Percentage In Decade
  FROM Unpivoted Sector Data
  GROUP BY Country_Code, (Year / 10) * 10, Sector_Name
Rank Sectors within Each Country-Decade to find the dominant one
Ranked Sector Dominance AS (
  SELECT
    Country_Code,
    Decade.
    Sector_Name,
    Avg_Sector_Percentage_In_Decade,
    ROW NUMBER() OVER(PARTITION BY Country Code, Decade ORDER BY Avg Sector Percentage In Decade DESC) as m
FROM Average Sector Contribution Per Decade
)
SELECT
  Country_Code,
  Decade,
  Sector_Name AS Dominant_Sector,
```

FROM Ranked Sector Dominance

WHERE m = 1

ORDER BY Country_Code, Decade;

OUTPUT:



→ Calculate CAGR for each Sector per Country

```
SELECT
```

Country_Code,

 $Sector_Name,$

First_Year,

Last_Year,

Beginning_Value,

Ending_Value,

(Last_Year - First_Year) AS NumberOfYears,

- -- CAGR Calculation: ((EndingValue / BeginningValue)^(1 / NumberOfYears)) 1
- -- Use POWER function for exponentiation. Handle division by zero or negative values.

CASE

```
WHEN (Last_Year - First_Year) > 0 AND Beginning_Value > 0 THEN
```

 $(POWER((Ending_Value \ / \ Beginning_Value), (1.0 \ / \ (Last_Year - First_Year))) - 1) * 100 -- Convert to percentage)) - 1) * 100 -- Convert to percentage) - 100 -- 1$

ELSE NULL -- Cannot calculate CAGR if duration is $\boldsymbol{0}$ or beginning value is $\boldsymbol{0}$

END AS CAGR_Sector_Pct

FROM First_Last_Sector_GDP

WHERE Beginning_Value IS NOT NULL AND Ending_Value IS NOT NULL

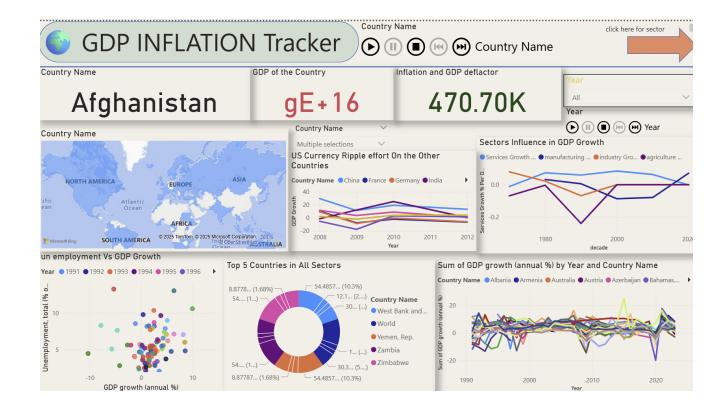
ORDER BY Country_Code, Sector_Name;

OUTPUT:

: Countr	Sector_N	First_Year	Last_Year	Beginning_Value	Ending_Value	NumberO
AUS	Agriculture					34
AUS						34
AUS	Manufactu			149128591012.68152	149128591012.68	34
AUS				1148449765590.7542	1148449765590.7	34
BRA	Agriculture			1082500052.8079178	1082500052.8079	43
BRA				4137840235.6691685	4137840235.6691	45
BRA	Manufactu			1561396848.7774792	1561396848.7774	45
BRA				11507519670.08321	11507519670.083	45
CHE		2002				20
CHE						20
CHE	Manufactu				437714814.99967	20
CHE				3371845489.7536006	3371845489.7536	20
CHN	Agriculture				22378887999.730	64
CHN	Industry	1960	2024	120467066195.7315	120467066195.73	⁶⁴ ₽ ≡

4. PowerBI Dashboard -1

Dashboard - 1 GDP INFLATION Tracker



Overview

The Power BI dashboard analyzes trends in GDP and inflation of Afghanistan while also showing a view of inflation effects (focusing on geography), GDP influences (by sector) and individual country performances.

Key Insights

Country Metrics (Afghanistan)

- GDP of the Country: 1.79E+16
- Inflation and GDP Deflator: 470.70K this is a high price level adjustment/inflation indexing over time

Geography and Global View

- World Map (Country Code Map): Shows the spatial extent and is selectable against country parameters globally
- US Ripple Effects on Other Countries:
- o This shows that the US Recession rippled outwards affecting countries along the way Japan, Russia, UK... etc.
- o The ripple effect is based upon corresponding GDP percentage changes of all affected countries.

Sector-wise GDP Influence over decades

- Influencing GDP Growth by sector (1975-2020) typically:
- o Services sector has generally always shown positive.
- o Manufacturing sector has some fluctuations while agriculture fluctuates with dipping below zero around years 1990's. o Industry shows that after years 2000, there has been recovery and stabilization.

GDP Growth by Year and Country

- The source is a clear comparison against countries:
- Afghanistan shows a quickly steep initial growth downward until it flattens,
- Through countries Africa Eastern & Southern secured relatively stable and low emergent GDP growth,
- The world averages are only inferred in this view and don't really require highlighting establishing any benchmarks.

GDP vs. Inflation

- Bubble Plot shows a weak or scattered correlation between GDP and inflation (consumer prices).
- A few countries show very high inflation and moderate GDP (outliers that may be worthwhile to investigate). **Top 5 Countries in All Sectors**

Donut chart includes:

• West Bank and Gaza, World, Yemen, Zambia, and Zimbabwe as top contributors in their sectors.

Their share in their sector contributions to global GDP (percentage and GDP amounts).

Issues of Note

- Service sector dominance is consistently across the globe across decades.
- US economic conditions matter the most to the changes in the GDP of allied or connected economies.
- Afghanistan's GDP appears extremely scaled and could possibly need unit validation.
- Inflation indicators such as GDP deflator or consumer price index, must be normalized, since they vary so wideley, for comparisons to be valid.

Us currency Ripple effect:

- You can observe how closely the established economies (**Germany, France, Japan**) track the US economy, likely showing a high degree of correlation during downturns.
- In contrast, you can see if the powerful growth trajectories of **China** and **India** show more independence, diverging from US trends except during major global shocks like the 2008 crisis.
- The comparison with **Mexico** would be particularly interesting to visualize the strong, interconnected trade relationship between the two neighbors.

Dashboard-2 - GDP Sector Dominance Tracker Agriculture Vs Industry



Overview of Dashboard

This dashboard compares the average sector contributions for Agriculture and Industry to GDP across decades and countries. It provides a way to visualize which sectors dominate over time, and documents the global and country-level ranks and shifts.

Key Findings

Leading Contributors by Sector

- For Industry, the countries still in the top-10 after each decade were:
- Saudi Arabia, Russian Federation, Korea (Rep.), Mexico, Switzerland, Türkiye, and Spain.
- Saudi Arabia ranks first in other decades with ~54.45% of their overall GDP being industrial contribution.
- The Top-10 agriculture contributions were (and specifically per decade):

 Türkiye, Korea (Rep.), Mexico, Russian Federation, Saudi Arabia, and Spain.
 Türkiye showed to be very consistent with agriculture contribution showing an average of 21.01%.

Country-level comparison

The table shows:

- Spain has the highest average agriculture contribution (2.97%) but ranked 15th for industry.
- Türkiye drops behind in the industry contribution (Rank 3 with avg agriculture contribution of 21.01% [Rank 1]) showing a balance of dominance:

- Agriculture avg.: 21.01% (Rank 3) Industry avg.: 26.01% (Rank 1)
- United States av. agriculture contribution (1.00%) with 20th place for industry.

Historical Sector Performance

The historical line graph shows overall contributions from agriculture and industry through time.

- It's clear industry has always been ahead of agriculture over time and especially after the year 1980 where visible growth arises.
- Agriculture is fairly stable with minor declines or flat lines over time. Dominant Sector by Decade

Bar chart shows sectoral dominance trend:

- Industry is the dominant sector in the majority of countries across all decades.
- Only during the early 1960s and 1970s, a few countries showed agriculture dominance, but the count steadily declines after 1980. o By 1990, industry dominance becomes overwhelming.

Agriculture vs Industry Rank

- Countries like India, China, and Brazil show higher agriculture rankings.
- **Developed nations** (e.g., US, UK, Germany) consistently rank higher in **industry**, confirming economic transition patterns.
- Emerging economies still maintain significant agriculture rankings.

Notable Observations

- Türkiye is the most balanced country with high ranks in both agriculture and industry.
- **Industrial dominance** has grown over the decades, marking structural shifts in global economies.
- **Developed countries** heavily industrialized; **developing countries** show more diversified or agriculture-leaning profiles.
- Some outliers (e.g., Saudi Arabia) show extreme dominance in industry due to resourcebased economies.

Dashboard-3 - GDP Sector Dominance Tracker Service Vs Manufacturing



Dashboard Purpose

This module provides data over time on the dominant sector between Services and Manufacturing for each country over time, allowing for long-term sectoral transitions and the ability to see which countries contribute the most.

Key Insights

Country View

- Selected Country: Afghanistan Overall Sector: Manufacturing
- This means that Afghanistan's economy contributes more towards industrial output than service output in its GDP.

Overall Global Sector Dominance

• Total Number of Countries: 252

Service Count: 1167

- In all decades and in all countries, the services sector seldomly appears with a manufacturing sector stronger.
- Nonetheless, this follows the overall economic trend globally of transitioning away from manufacturing to a service-oriented economy, even more so in developed and emerging economies.

Decadal Trends: Service vs Manufacturing •

Sum of In Contributions for Years 1960-2025:

- Both sectors have long-term growth, however Services sector has seen an upward trajectory since the 1990s.
- The manufacturing sector saw its peak around the 2010s, but has either remained stagnant or declined since.
- Growth % change per decade:
- Services Growth % change was at its lowest growth prior to 2010, however following 2010 it has risen significantly and eclipsed that of manufacturing sector change post2020.
- Manufacturing Growth % change was at the lowest recorded between 1990s-2000s and likely linked to the transition of economies towards a service provided structure following the advent of globalization.
- It appears the quarterly commentary from 2000-present has seen a very slight increase in growth and not all linked back to the service transition but rather industrial policy upliftings in many developing economies. Top Performing Countries
- Top 10 Services Countries (by GDP contribution):
- Antigua and Barbuda, Argentina, Arab World, Africa Eastern/Southern, and Afghanistan (lower but included).
- Indicates smaller or tourism-heavy economies can have GDP levels with High service ratios.
- Top 10 Manufacturing Countries:
- Zimbabwe, Zambia, Uruguay, Uzbekistan, Vietnam, Yemen, and others.
- Strong representation of developing nations that have led with industry.

Interesting Observations.

- Afghanistan likely has a troubling eventuality of No manufacturing which should be variable, as Manufacturing, like the developing economies that transitioned towards services to 2020.
- Services is now the preferred, most dominant sector globally to have emerged especially since 2000.
- While several African and Latin countries continue to show High manufacturing presence, they are undergoing industrial development.PowerBI Wireframe

PowerBI Wireframes:

1. Dashboard: GDP Inflation Tracker & Global View Placement Element Type Purpose / Title

Placement	Element Type	Purpose / Title
Top Bar	Text Box	Main Title: "GDP INFLATION TRACKER"
Top Bar	Slicer	Filter by "Country Name"
Top Bar	Slicer	Filter by "Decade"
Top Center	Text Box	Displays the Selected Country Name
Top Center	KPI Card	"GDP of the Country"
Top Center	KPI Card	"Inflation and GDP Deflator"
Left Panel	Map	Geographic location of the selected country
Left Panel	Scatter Plot	Correlation of "Sum of Inflation" vs. "Sum of GDP %"
Center	Line Chart	"US Ripple Effect on Other Countries"
Right Panel	Line Chart	"Sectors Influence in GDP Growth" (All four sectors)
Bottom	Donut Chart	"Top 5 Countries in All Sectors"
Bottom	Line Chart	"GDP Growth % by Year and Country Name"

2. Dashboard: Agriculture Vs Industry

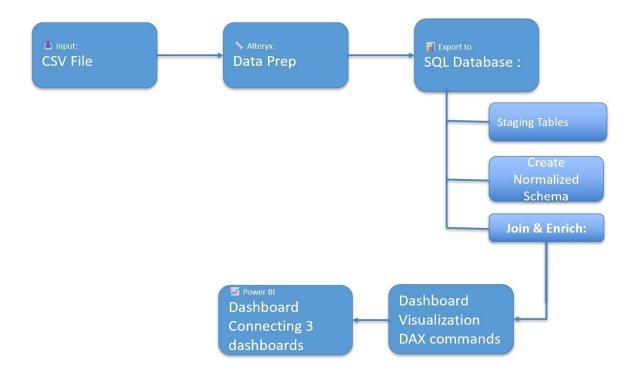
Placement	Element Type	Purpose / Title
Top Bar	Text Box	Main Title: "GDP Sector Dominance Tracker"
Top Bar	Slicer	Filter by "Select Country"
Top Bar Top Bar	Slicer Button	Filter by "Decade" Navigation to "Next page"
Left Panel	Donut Chart	"TOP 10 Industry per Decade"
Left Panel	Line Chart	"Agriculture vs Industry Rank"
Center Placement	Map Element Type	Geographic location of the selected country Purpose / Title

Center	Stacked Area Chart	"Historical Analysis" of Agriculture vs. Industry
Right Panel	Table	Detailed comparison of average sector values and ranks
Right Panel	Donut Chart	"TOP 10 Agriculture Per Decade"
Right Panel	Column Chart	Count of "Dominant Sector" by Decade (Agri vs Industry)

3.Dashboard: Service Vs Manufacturing

Placement	Element Type	Purpose / Title
Top Bar	Text Box	Main Title: "GDP Sector Dominance Tracker"
Top Bar	KPI Card	"Service Dominance Count"
Top Bar	KPI Card	Total "Countries" Tracked
Left Panel	Slicer	Filter by "Country Name"
Left Panel	Slicer	Filter by "Decade"
Left Panel	Text Box	Displays selected country & dominant sector (e.g., "Afghanistan Manufacturing")
Left Panel	Map	Geographic location of the selected country
Left Panel	Bar Chart	"TOP 10 Service Countries"
Right Panel	Stacked Area Chart	"Service vs Manufacturing" Sum over Time
Right Panel	Line Chart	"Service vs Manufacturing over Decade" (Growth %)
Right Panel	Bar Chart	"TOP 10 Manufacturing Countries"

5. DATA PIPELINE DIAGRAM



Githublink: https://github.com/guruc267/GDP-Sector-Dominance-Tracker-