DATA WAREHOUSE SOLUTIONS: ENHANCING VIRGINIA'S REGIONAL GDP ANALYTICS

Course: Database Management

By: Parth Parker

MSBA

CONTENTS

DATA & PROBLEM INTRODUCTION

Business Description

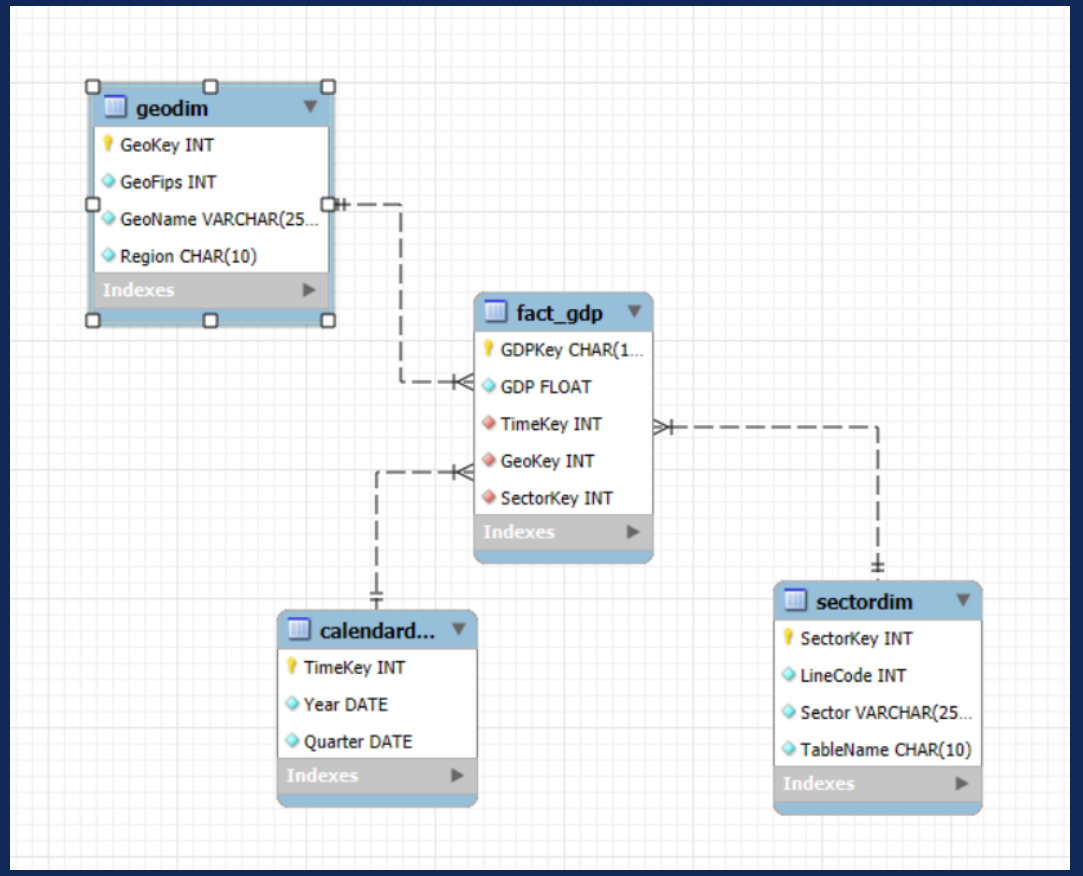
- Overview: Dataset on GDP by year and quarter for the state of Virginia.
- To help analyze sector wise, quarter wise, geography wise GDP per cap for the state of Virginia, to allocate investments/resources optimally.
- Variables: "_id", "Year", "Quarter", "GeoFips", "GeoName", "Region", "TableName", "LineCode", "Description", "Unit", "DataValue".

Information & Source Data

- Virginia Gross Domestic Product Quarterly Data
  - Provides information on the active economic sectors in VA and their impact on the states GDP.
  - Data was sourced from the Virginia Open Data Portal.

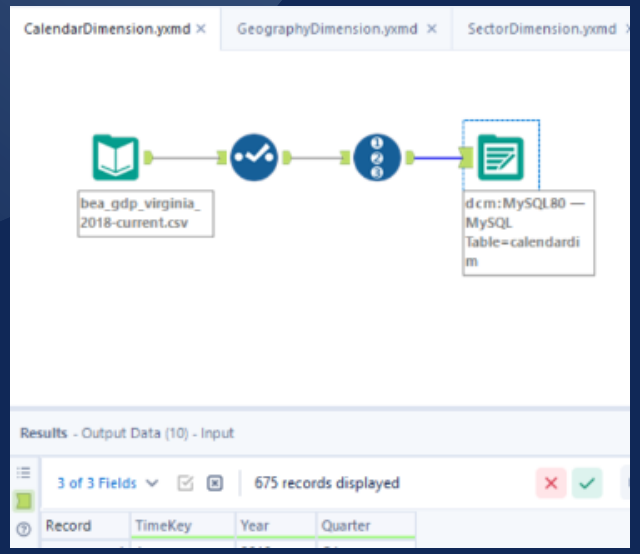
DIMENSIONAL MODEL DIAGRAM

- **Calendar Dimension**: For tracking time periods. Year and quarter.
- **Geography Dimension**: For location-specific data. Only Virginia.
- **Sector Dimension**: To categorize GDP by sectors.
- **Fact Table**: Contains GDP values per sector without any additional economic indicators.



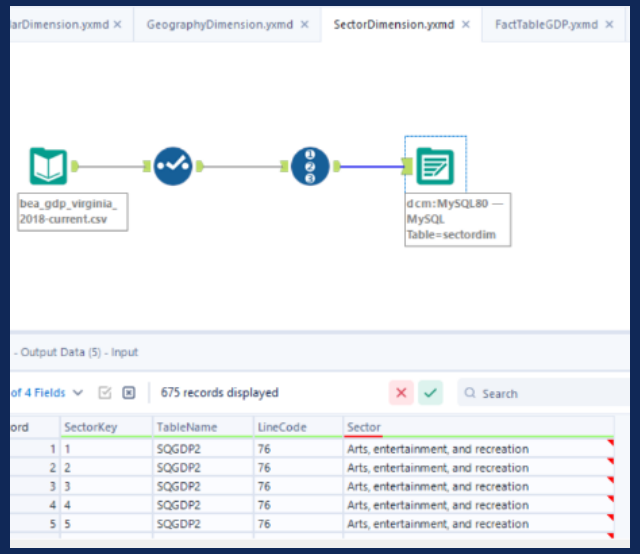
ETL PROCESSES

Geography Dimension



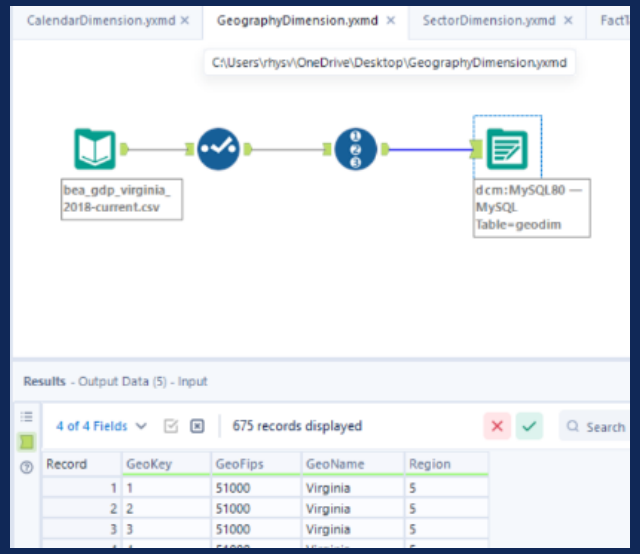
Joining: Linked dimensions (Calendar, Geography, Sector) to the Fact Table.

Sector Dimension



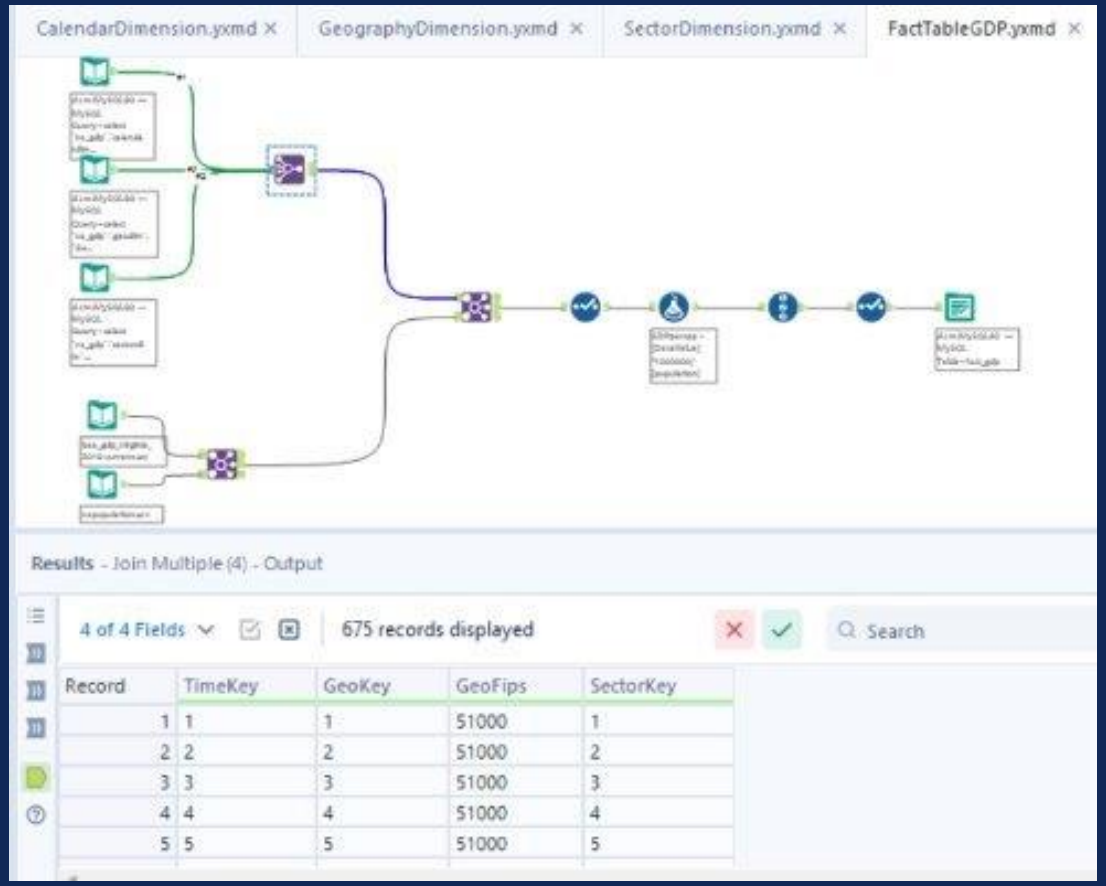
Calculations: Computed GDP per capita.

Calendar Dimension



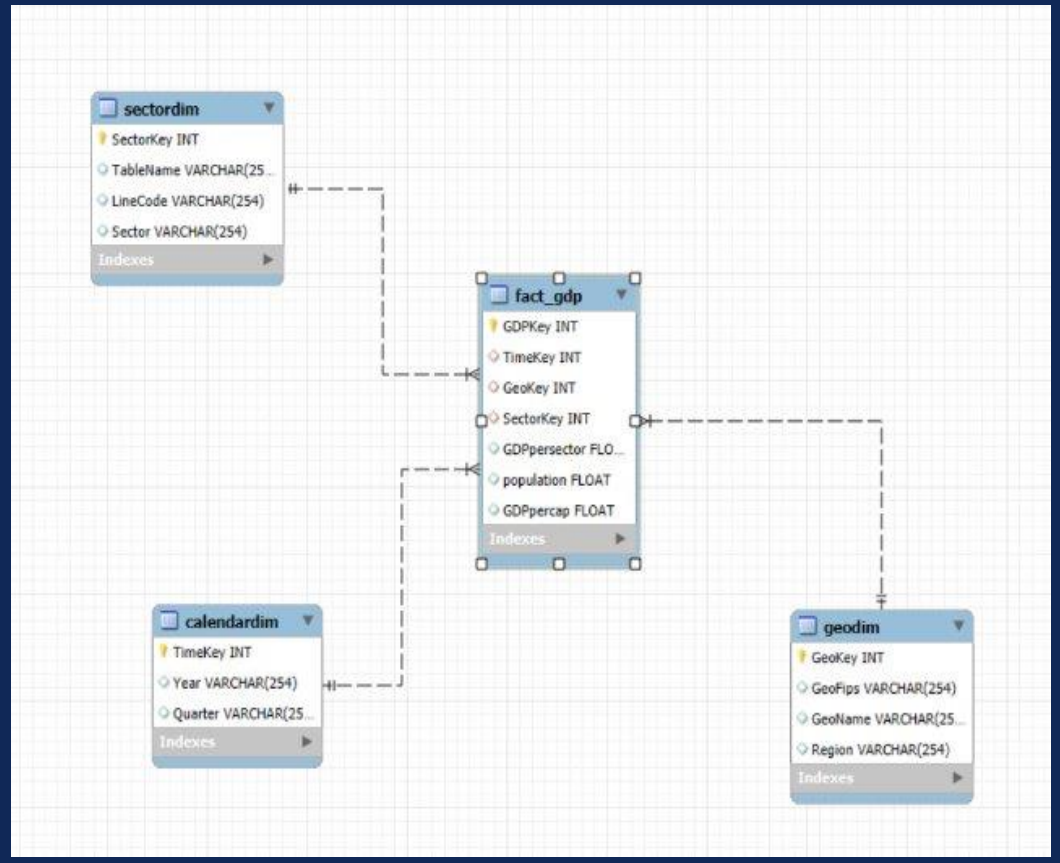
Data Cleaning: Standardized formats and removed duplicates.

FACT TABLE ETL PROCESS



- Integrates the Calendar, Geography, and Sector dimensions
- Uses *JOIN MULTIPLE* tool to link records by *TIMEKEY*, *GEOKEY*, and *SECTORKEY*
- Calculates GDP per capita using Virginia's population data for enriched analysis
- Outputs the final table to MySQL, and is ready for multi-dimensional analysis by time, location, and sector

FINAL SCHEMA

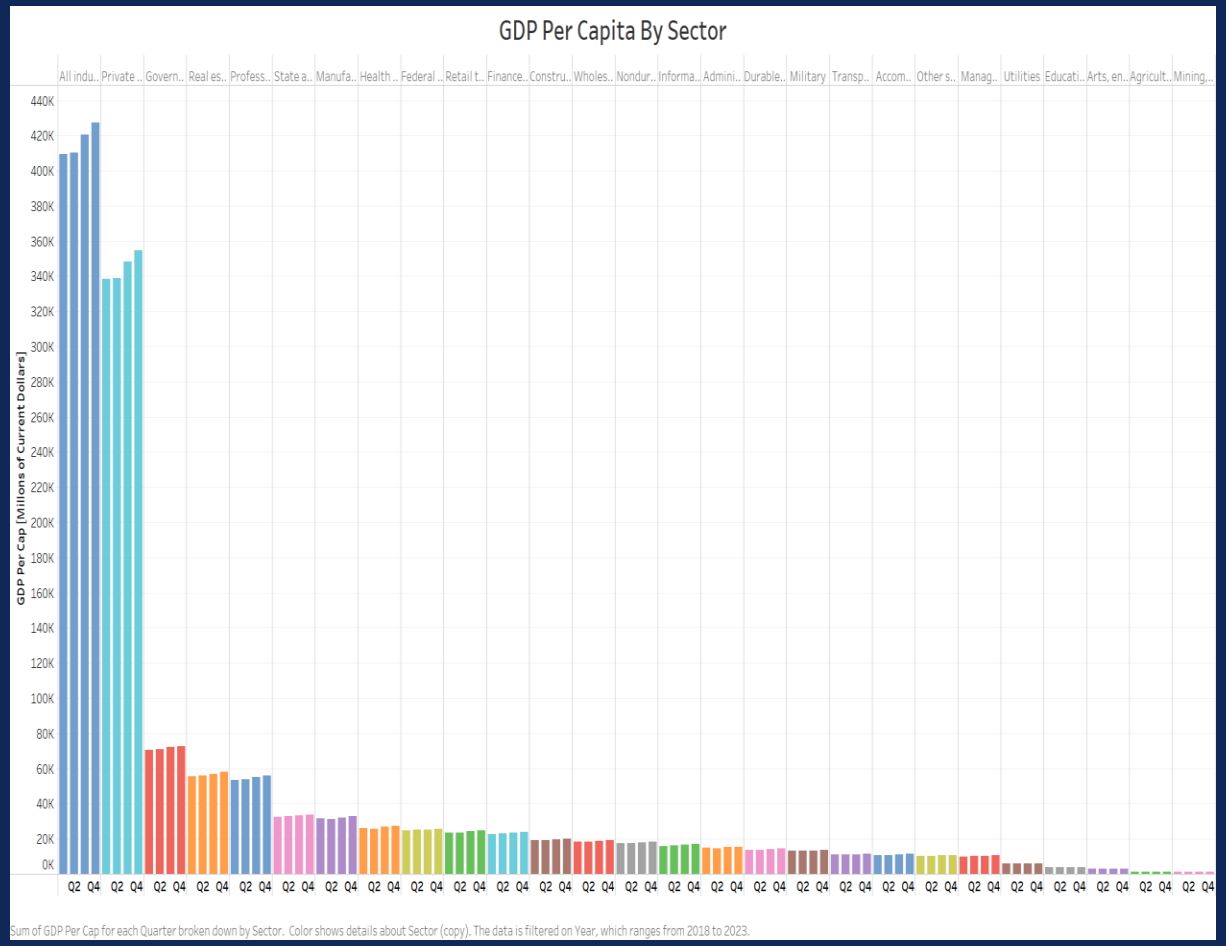


- **Key Schema Enhancements:**
  - Population Data: Allow for GDP per capita calculation.
  - New Fact Table Structure: The fact table now includes fields for GDP per capita, population, and total GDP.
  - ETL Transformation: Joined multiple datasets to create a cohesive model.

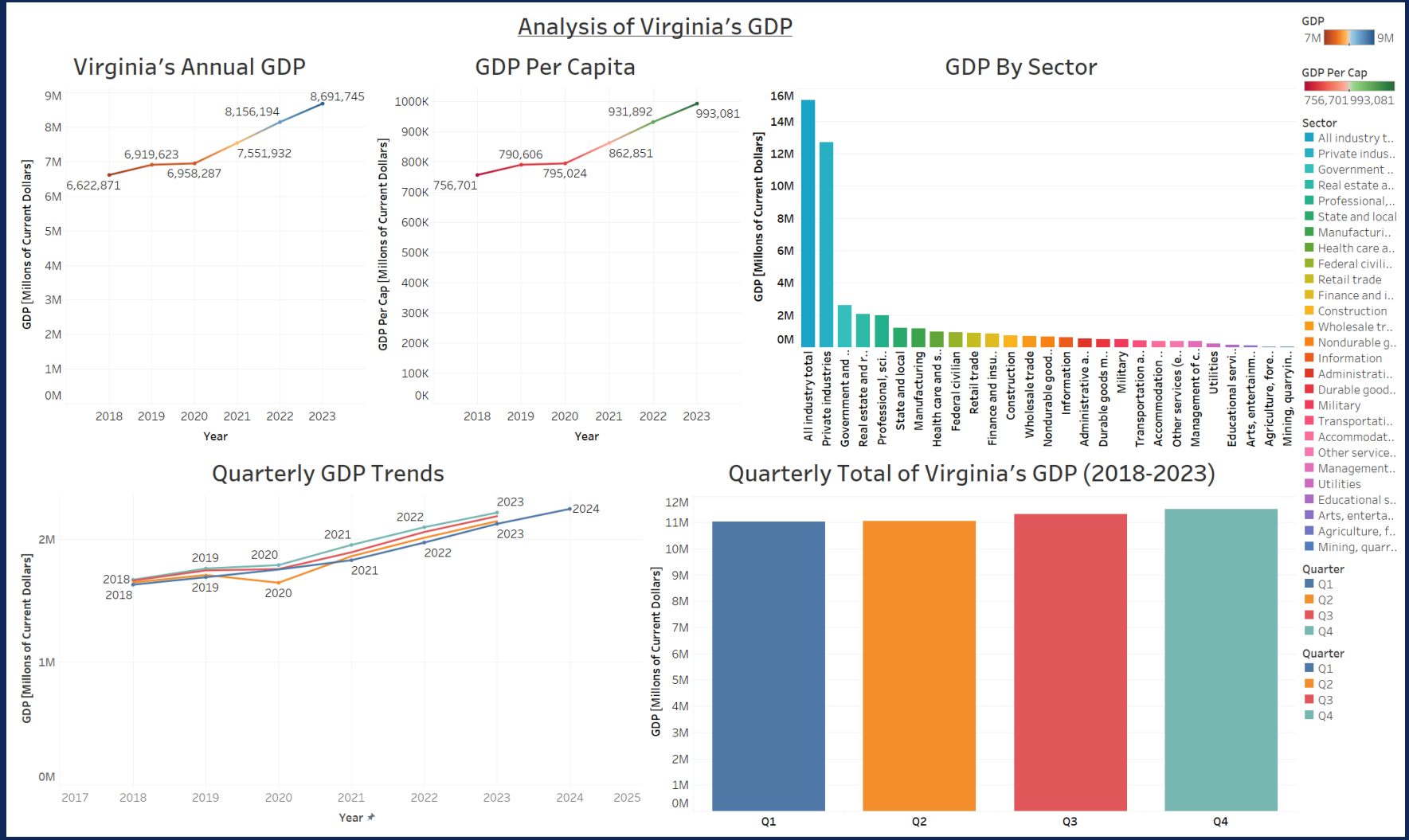
ANALYSIS

What's Changed?

- **Capabilities**: Able to analyze GDP by time values, geography, and sector, for more in-depth insights about Virginia's GDP
- **Key Metrics**: Total GDP per sector and GDP per capita for welfare analysis
- **Types of Analysis**: We can now perform multiple types of analysis including
  - Trend
  - Regional/National comparisons
  - Time Series
  - Sector
- **Future Insights**: expandable with other metrics such as employment rates, updated population, CPI, and numerous other factors that affect or indicate the direction of the Virginia's GDP



ANALYSIS


Analysis of Virginia's GDP
Virginia's Annual GDP
GDP [Millions of Current Dollars]
6,622,871
6,919,623
6,958,287
8,156,194
7,551,932
8,691,745
Year
GDP Per Capita
GDP Per Cap [Millions of Current Dollars]
756,701
790,606
795,024
931,892
862,851
993,081
Year
GDP By Sector
GDP [Millions of Current Dollars]
All industry total
Private industries
Government and ..
Real estate and r..
Professional, sci..
State and local
Manufacturing
Health care and s..
Federal civilian
Retail trade
Finance and insu..
Construction
Wholesale trade
Nondurable good..
Information
Administrative a..
Durable goods m..
Military
Transportation a..
Accommodation ..
Other services (e..
Management of c..
Utilities
Educational servi..
Arts, entertainm..
Agriculture, fore..
Mining, quarryin..
GDP
7M
9M
GDP Per Cap
756,701 993,081
Sector
All industry t..
Private indus..
Government ..
Real estate a..
Professional,..
State and local
Manufacturi..
Health care a..
Federal civili..
Retail trade
Finance and i..
Construction
Wholesale tr..
Nondurable g..
Information
Administrati..
Durable good..
Military
Transportati..
Accommodat..
Other service..
Management..
Utilities
Educational s..
Arts, enterta..
Agriculture, f..
Mining, quarr..
Quarter
Q1
Q2
Q3
Q4
Quarterly GDP Trends
GDP [Millions of Current Dollars]
Year
Quarterly Total of Virginia's GDP (2018-2023)
GDP [Millions of Current Dollars]
Q1
Q2
Q3
Q4

THOUGHTS AND CONCLUSIONS

- Value of the Model: Structured data enables detailed GDP analysis.
- Population Data Insight: GDP per capita adds depth to sector and region comparisons.
- Learning: ETL challenges emphasized data consistency and accuracy.
- Future Potential: Easily expandable for broader economic studies, with different factors and values
- Summary: Dimensional modeling supports data-driven insights and decision-making. Making it easier to connect important values together and perform more thoughtful analysis

FINAL INSIGHTS & REFLECTIONS

System Realization

- Creating our data warehouse enforces consistency across different data sources.
- Ensures reliability when running queries and analytics.
- Easily add other metrics to our dataset using Alteryx. Such as unemployment data, or GDP data prior to 2018.
- Ultimately, we made the dataset more valuable to analysts.

Team Thoughts

- Parth - The future scopes excite me the most, with joining data from multiple states including socioeconomic factors
- Rhys - I found the joining the fact table the most challenging. While that's the case, it's the part I most enjoyed.
- Brendon – I enjoy presenting the findings and being able to tell a story about the business.
- Shields – I liked the data analysis and being able to show how changes are made in both in the business and the factors that affect them

THANK YOU FOR LISTENING

Are there any questions?