Project Development Phase

Data Preprocessing Steps

1. Data Collection

- Downloaded the **Indian Union Budget dataset (FY 2021–2024)** in CSV format.

- Verified dataset completeness and structure.

- Ensured attributes such as ministry, scheme, revenue, capital, and total budget are available.

2. Understand the Dataset

- Identified column names and their meanings.

- Checked data types (numeric, text).

- Identified missing values and inconsistencies.

- Reviewed business requirements related to healthcare funding analysis and budget allocation trends.

3. Import Dataset into Database / Tableau

- Loaded dataset into **MySQL / Tableau**.

- Validated successful import without errors.

- Ensured no data truncation or formatting issues.

4. Connect Tableau Desktop to Database Server

- Established connection between Tableau and MySQL data source.

- Verified tables and fields are visible in Tableau.

- Ensured data was accessible for visualization.

5. Data Preparation

Remove Duplicate Records

- Checked and removed duplicate entries if found.

Handle Missing Values

- Verified null values in budget fields.

- Replaced or handled missing values where required.

Correct Data Types

- Converted budget columns to numeric values.

- Ensured ministry and scheme fields are text.

- Verified year-wise budget fields are correctly formatted.

Data Standardization

- Standardized budget units and formats.
- Ensured consistent naming for ministries and schemes.
- Formatted financial values properly for analysis.

Create Calculated Fields in Tableau

- Total Budget Calculation
- Year-wise Budget Comparison
- Percentage Contribution of Ministry
- Healthcare Funding KPI metrics

Filtering & Cleaning

- Removed unnecessary columns.
- Filtered data based on healthcare-related analysis requirements.

6. Data Validation

- Cross-checked totals and aggregations.
- Ensured no null values affecting dashboards.
- Verified visualization outputs match dataset values.

Business Questions with Visualizations

1. Which ministry receives the highest healthcare allocation?

Visualization: Bar Chart

- X-axis: Ministry / Department
- Y-axis: Total Budget
- Sorted descending

Purpose: Identify primary funding authorities influencing healthcare investment.

2. Which healthcare schemes receive the maximum funding?

Visualization: Horizontal Bar Chart

- Y-axis: Scheme
- X-axis: Total Budget
- Top 5 filter applied

Purpose: Identify priority programs for potential collaboration or alignment.

3. How has healthcare funding changed from 2021–2024?

Visualization: Line Chart

- X-axis: Financial Year
- Y-axis: Total Budget

Purpose: Analyze funding growth trends and investment direction.

4. What is the distribution between Revenue and Capital expenditure?

Visualization: Donut Chart or Pie Chart

- Dimension: Budget Type (Revenue / Capital)
- Measure: Total Allocation

Purpose: Understand whether spending is operational or infrastructure-focused.