Preprocessing and Business Questions Report

1. Data Preprocessing Steps Applied

1. Standardized column names (converted to lowercase and replaced spaces with underscores).
2. Removed duplicate records to ensure data integrity.
3. Handled missing values:
4. Numerical columns filled using median values.
5. Categorical columns filled using mode values.
6. Converted numerical columns to appropriate numeric data types.
7. Detected and removed outliers using the Interquartile Range (IQR) method.
8. Performed final null-value validation and cleaned residual missing entries.
9. Exported the final processed dataset as Cleaned_Dataset.csv for visualization in Tableau.

2. Key Business Questions Addressed

1. Which departments received the highest budget allocations in each financial year?
2. How has total budget allocation evolved from FY 2021–22 to FY 2023–24?
3. Which sectors show consistent growth or decline in allocation trends?
4. What percentage of the total budget is allocated to top ministries?
5. How does capital expenditure compare to revenue expenditure across years?
6. Which categories demonstrate the highest investment impact?
7. Are there noticeable shifts in allocation priorities over the analyzed years?

3. Conclusion

The preprocessing phase ensured data accuracy, reliability, and readiness for analytical visualization.

The cleaned dataset enabled meaningful insights through Tableau dashboards and stories, supporting data-driven analysis of Union Budget allocation trends and sustainable growth strategies.