Data Collection and Preprocessing Phase

Date	7 March 2025
Team ID	Prakriti Gupta (Individual Project)
Project Title	Measuring the pulse of prosperity :An Index Of Economic Freedom Analysis
Maximum Marks	10 Marks

Data Exploration and Preprocessing Template

Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

Section	Description
Data Overview	The dataset contains information on the **Economic Freedom Index (2022)**, including country-wise scores, rankings, GDP growth rates, inflation rates, unemployment rates, financial freedom scores, and other macroeconomic indicators. The dataset was sourced from **SmartInternz workspace**.
Data Cleaning	**Handled Missing Values:** Used mean/median imputation for numerical fields and mode for categorical ones. Dropped rows with critical missing data. **Removed Duplicates:** Eliminated duplicate records based on country names, ranks, and scores. **Corrected Errors:** Standardized country names, fixed negative GDP growth inconsistencies, and verified outliers in Economic Freedom Scores. **Formatted Data:** Ensured proper numerical formatting for accurate calculations.

Data Transformation	● **Filtering & Sorting:** Extracted the top 40 ranked countries and sorted them by Economic Freedom Score. ● **Pivoting:** Restructured data to compare rankings across different indicators. ● **Calculated Fields:** Created new metrics like GDP per capita growth rate and inflation-adjusted scores for deeper insights.
Data Type Conversion	● Converted numerical fields like GDP, inflation rate, and index scores to appropriate data types (float/integer). ● Standardized date formats where applicable. ● Ensured categorical fields like country names remained consistent for accurate filtering and visualization.
Column Splitting and Merging	● Split combined fields (e.g., country and region) for better granularity. ● Merged relevant attributes to create new insights, such as combining financial freedom and inflation data for correlation analysis.
Data Modeling	● Established relationships between tables using country names as primary keys. ● Ensured proper joins for accurate aggregation in Tableau. ● Optimized data structure for efficient querying and visualization.
Save Processed Data	● Exported the cleaned and transformed data as a CSV for future analysis. ● Stored processed data in MySQL for easy retrieval and integration with Tableau. ● Ensured data integrity by maintaining backups of raw and cleaned datasets.