Data Collection and Preprocessing Phase

Date	01 October 2025
Team ID	SWUID20250181698
Project Title	Global Inflation Dynamics (1980-2024): A Comparative Time-Series 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 inflation data for 196 countries across 45 years (1980-2024). It measures the Annual Average Inflation (Consumer Prices) rate.
Data Cleaning	**1. Missing Values:** Rows with null Inflation_Rate were dropped. **2. Country Filtering:** Countries with fewer than 23 years (50% of the 45-year range) of non-missing data were removed using a Group By function in Power Query. **3. Duplicate Handling:** No duplicates were found at the (country_name, Year) level after unpivoting.
Data Transformation	The core transformation used Power Query's **Unpivot Other Columns** feature. This converted the data from a **Wide Format** (47 columns) to a **Long Format** (approx. 7,952 rows), which is essential for time-series plotting.
Data Type Conversion	The Year column (originally Attribute) was converted to a Whole Number, and the Inflation_Rate column (Value) was converted to a Decimal Number.

Column Splitting and Merging	A new column, **Bloc**, was created using Power Query's **Conditional Column** feature to categorize countries as 'G7 (Developed)', 'BRICS (Emerging)', or 'Other', which is crucial for the comparative analysis visualization.
Data Modeling	**Measures Creation (DAX):** Since this is a single table, the primary modeling task is defining the **DAX measures**, specifically the **Median Inflation** measure, to handle outliers.
Save Processed Data	The processed and cleaned table is loaded from Power Query into the Power BI data model for visualization.