**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 21 June 2025 |
| Project Title | Power BI Inflation Analysis – Journeying Through Global Economic Terrain |
| Maximum Marks | 10 Marks |

**Data Exploration and Preprocessing**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | The dataset was sourced from Kaggle and includes country-wise and global inflation data (primarily Consumer Price Index – CPI) from 1960 to 2023 in CSV/Excel format. Each file contains inflation values mapped to countries and corresponding years. |
| Data Cleaning | Null values and blank records were removed. Duplicates were eliminated to ensure dataset uniqueness. Missing inflation values were imputed using median values or forward-fill methods. Outliers were identified but retained for economic pattern visibility. |
| Data Transformation | Power Query was used for cleaning and transforming the data. Inflation values were filtered by year and country, sorted, and normalized. Tables were unpivoted to follow a tidy data structure with “Country,” “Year,” and “Inflation (%)” as columns. |
| Data Type Conversion | All numeric columns such as inflation rates were converted to float types. Year was set as an integer, and country names were standardized as categorical data. Text encodings and date formats were corrected for consistency. |
| Column Splitting and Merging | No column splitting was required. However, inflation rates from multiple files were merged into a consolidated master table using Power BI's data model features. The result was a uniform table with "Country," "Year," and "Inflation" columns. |
| Data Modeling | Data modeling in Power BI established relationships for cross-country and time-series analysis. Measures like Global Average Inflation, Country Max/Min Inflation, and 5-year Average CPI were created using DAX functions (e.g., AVERAGEX, MAXX). |
| Save Processed Data | The final dataset was saved in .pbix format for Power BI. Backup CSVs were also stored for traceability. All preprocessing steps, transformation logic, and calculated measures were documented for reproducibility and audit readiness. |