**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 02 Oct 2025 |
| Name | Muhammad Yasir |
| Team ID | xxxxxx |
| Project Title | **Power BI Inflation Analysis: Journeying Through Global Economic Terrain** |
| Maximum Marks | 2 Marks |

## ****Data Collection Plan****

### ****Project Overview****

The project focuses on analyzing global inflation trends across countries from 1980 to 2024. The primary objective is to identify inflation patterns, regional differences, missing data gaps, and provide a clean, standardized dataset suitable for creating dashboards and conducting advanced analysis in Power BI and Excel.

### ****Data Collection Plan****

Data will be collected from:

* **International financial institutions** (World Bank, IMF, OECD, UN Data).
* **Official government statistical agencies** (for country-specific inflation records where available).
* **Third-party open data repositories** (Kaggle, data.world, and GitHub).
* **Manual enrichment** using ISO country codes and regional classifications (e.g., UN geoscheme, World Bank regional groupings).

### ****Raw Data Sources Identified****

* **Primary Inflation Dataset (1980–2024)**: Covers annual inflation rates for 180+ countries.
* **Country-Region Mapping Dataset**: Provides region and subregion classification for each country.
* **Supplementary Data** : Hyperinflation case studies (Venezuela, Zimbabwe, Lebanon), missing years backfilled using IMF/World Bank datasets.

## ****Raw Data Sources****

| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| --- | --- | --- | --- | --- | --- |
| Inflation Dataset (1980–2024) | Annual inflation rates by country from 1980–2024. Includes missing values and outliers. | Provided dataset (global\_inflation\_data.csv) | CSV | ~5 MB | Public (project dataset) |
| World Bank – World Development Indicators | Contains macroeconomic indicators including inflation, CPI, and GDP. | <https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG> | CSV/Excel | Varies | Public |
| IMF Data Portal | Inflation and macroeconomic data (supplementary for missing values). | <https://www.imf.org/en/Data> | Excel/CSV | Varies | Public |
| Country-Region Mapping (UN Geoscheme) | Country classification into region/subregion (Asia, Europe, Africa, etc.). | <https://unstats.un.org/unsd/methodology/m49/> | CSV | Small | Public |
| Kaggle Open Datasets | Community-provided inflation datasets for cross-validation. | <https://www.kaggle.com/> | CSV/Excel | Varies | Public (requires login) |