**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 01 October |
| Team ID | SWUID20250181698 |
| Project Title | Global Inflation Dynamics (1980-2024): A Comparative Time-Series Analysis |
| Maximum Marks | 3 Marks |

**Data Quality Report Template**

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| global\_inflation\_data.csv | Structural Issue: Wide Format (Years are columns, not rows). | High | **Transformation:** Use Power Query's **Unpivot Columns** function to convert years (1980-2024) into a single Year column and a single Inflation\_Rate column |
| global\_inflation\_data.csv | Missing Values (NaNs): Numerous null values, particularly for earlier years or small countries. | Moderate | Filtering: Use Power Query to Remove Empty rows in the new Inflation\_Rate column. Additionally, filter out countries with less than 50% of data points for reliability. |
| global\_inflation\_data.csv | Data Type Mismatch: Year columns are imported as Text instead of a number, and inflation rates need to be treated as a decimal number. | Low | Type Conversion: Explicitly convert the unpivoted Year column to Whole Number and the Inflation\_Rate column to Decimal Number. |
| global\_inflation\_data.csv | Outliers / Hyperinflation: Extreme values (e.g., Venezuela's rates reaching thousands of percent) that would skew the average. | Low | Mitigation via DAX: Do not use the AVERAGE function in Power BI. Instead, create a DAX Measure using the MEDIAN function for all group aggregations to ensure the central tendency is robust. |