**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | 740012 |
| Project Title | **Predicting IMF-Based Exchange Rates: Leveraging Economic Indicators for Accurate Regression Modeling** |
| Maximum Marks | 2 Marks |

**Data Quality Report:**

The Data Quality Report 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 Quality Report:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| Kaggle  Dataset | 1. Missing values: Gaps in economic indicator data or exchange rate data.  2. Outliers: Erroneous or unusual data points that can skew the model.  3. Noise: Random fluctuations in the data that can affect the model's accuracy.  4. Inconsistent data: Differences in data formats, frequencies, or units.  5. Data granularity: Insufficient detail in the data, making it difficult to capture important patterns. | Moderate | Use mean/median imputation. |
| Kaggle  Dataset | Categorical data in the dataset | Moderate | Encoding has to be done in the data. |