**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 20 July 2024 |
| Team ID | 739860 |
| Project Title | A Comprehensive Measure Of Well-Being : A  Human Development Index Using Machine Learning |
| 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:**

|  |  |  |  |
| --- | --- | --- | --- |
| Kaggle  Dataset | Categorical data in the dataset | Moderate | Encoding has to be done in the data(Labelencoding). |

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| Kaggle  Dataset | Missing values in the Dataset columns. | Moderate | Use mean/median imputation. |