**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 20 June 2024 |
| Team ID | 740026 |
| Project Title | Optimizing Sleep Efficiency: Harnessing Machine Learning For Enhanced Restorative Rest |
| Maximum Marks | 2 Marks |

**Data Quality Report Templat**

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** |
| Kaggle Dataset | Missing values in the  ‘Awakenings’, ’Caffeine consumption’, and ‘Alcohol consumption’ columns. | Low | Use Mean and Median/Mode imputation |
| Kaggle Dataset | Categorical data in the dataset | Low | Using LabelEncoder |