 

Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 19 April 2024 |
| Team ID | Team-738178 |
| Project Title | Envisioning Success : Predicting University Scores With 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:

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| Kaggle Dataset | Missing values in the 'num\_students', ‘student\_staff\_ratio’, ‘international\_students’, ‘female\_male\_ratio, columns. | Moderate | Use mean/mode  imputation. |
| Kaggle Dataset | Columns with Incorrect  data type ‘world\_rank’,  ’female\_male\_ratio’,’income’,  ’total\_score’,  ’international\_students’,  ’international’ | Moderate | Modify the data type of columns in the data. |