Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 18-06-2025 |
| Team ID | SWTID1749841176 |
| Project Title | Online Payments Fraud Detection 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:

|  |  |  |  |
| --- | --- | --- | --- |
| Data Source | Data Quality Issue | Severity | Resolution Plan |
| Kaggle Dataset | Values larger than float64 found in ‘amount’ column. | Moderate | Covert that data to null and then fill the null value with the median of that attribute. |
| Kaggle Dataset | Categorical data in the dataset | Moderate | Encoding has to be done in the data. |