**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 9 July 2024 |
| Team ID | SWTID1720019244 |
| Project Title | Traffictelligence: Advanced Traffic Volume Estimation with Machine Learning |
| Maximum Marks | 2 Marks |

**Data Quality Report Template**

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 ‘holiday’, ’temp’, ‘rain’ , ’snow’ and ‘weather’ columns | Moderate | Use mean/median and also frequently used sting imputations. |
| Kaggle Dataset | Categorical data in the dataset | Moderate | Encoding has to be done in the data. |