**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 13 July 2024 |
| Team ID | 740073 |
| Project Title | Exploratory Analysis of Rain Fall Data in India for Agriculture |
| 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 Data”**  **“Outliers”**  **“Inconsistent Data”**  **“Spatial Resolution”**  **“Temporal Resolution”**  **“Measurement Errors”**  **“Data Integration Issues”**  **“Historical Data Quality”**  **“Geographical Coverage”** | Moderate | Use mean/mode imputation. |
| Kaggle  Dataset | Categorical data in the dataset | Moderate | Encoding has to be done in the data. |