** **

**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 24 April 2024 |
| Team ID | 740140 |
| Project Title | Crystal Ball Analysis: Projecting Share Prices Of The Leading Gpu Titans |
| 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 price data for Q1 2023 | | High | Impute missing values with historical average |
| Kaggle Data | Inconsistent currency formats | | |  | | --- | | Medium |  |  | | --- | |  | | Standardize currency to USD  test\_data.isna().sum()  train\_data.isna().sum() |
| Kaggle Dataset | | |  | | --- | | Outliers in trading volume data |  |  | | --- | |  | | Low | Remove outliers using statistical methods |