

Data Collection and Preprocessing Phase

| | |
|---------------|--|
| Date | 9 th July 2024 |
| Team ID | SWTID1720435231 |
| Project Title | Walmart Sales Analysis For Retail Industry With Machine Learning |
| Maximum Marks | 2 Marks |

Data Quality Report:

The Walmart Sales Analysis source's data quality problems, in addition to their degrees of severity and proposed solutions, will be summarised in the Data Quality Report.

Data Quality Report:

| Data Source | Data Quality Issue | Severity | Resolution Plan |
|----------------|--|----------|---|
| Kaggle Dataset | Null values present in 'Markdown1', 'Markdown2', 'Markdown3', 'Markdown4', 'Markdown5' | Moderate | The given Null values present are replaced with zeros |
| Kaggle Dataset | Negative Values present in 'Weekly_sales' | Low | The negative values were omitted and only the non-negative data were taken into consideration |
| Kaggle Dataset | The dataset contains categorical data | Moderate | Encoding has to be done in the data. |