**Data Collection and Preprocessing Phase**

| Date | June 2024 |
| --- | --- |
| Team ID | 740053 |
| Project Title | Estimating the Stock keeping units using Machine Learning |
| Maximum Marks | 2 Marks |

**Data Collection Plan & Raw Data Sources Identification Template**

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data duration and integrity for informed decision-making in every analysis and decision-making endeavor.

**Data Collection Plan Template**

| **Section** | **Description** |
| --- | --- |
| Project Overview | Aims to provide accurate prediction estimates by considering factors like stock,total price, and promotions , validate its performance on unseen data to ensure it generalizes well. |
| Data Collection Plan | Data will be collected from order databases, stock prediction reports, demand forecasts, GPS data, and historical sales data. |
| Raw Data Sources Identified | Gathering data from order histories, delivery logs, prediction reports, demand forecasts, and ESUs integrations. |

**Raw Data Sources Template**

| **Source Name** | **Description** | **Location/URL** | **Format** | **Access Permissions** |
| --- | --- | --- | --- | --- |
| Dataset 1 | In this project we have used .csv data. This data is downloaded from kaggle.com. | https://www.kaggle.com/datasets/aswathrao/demand-forecasting | CSV | Public |