Data Collection and Preprocessing Phase

Date	10th July 2024
Team ID	SWTID1720086535
Project Title	E-commerce Shipping Prediction 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 curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	commerce shipping prediction estimates if a product will arrive on time, considering origin, destination, shipping method, carrier, and potential delays. Machine learning models use historical data and real-time updates, factoring in weather, traffic, and other variables. Accurate predictions enhance delivery estimates and customer experience, making it crucial for e-commerce businesses.
Data Collection Plan	Search for datasets related to e-commerce, shipping information, and customer details.
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle, a popular platform for data science competitions and repositories. The provided sample data represents a subset of the

	collected information, encompassing variables such as warehouse, product cost, customer ratings for machine learning analysis.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	. The "Customer Analytics" dataset consists of various features related to customers' demographics and behavior. It includes detailed information on customers' age, gender, Product cost, Warehouse details, shipment method, and more. This data is crucial for businesses looking to perform in-depth customer analysis and predict if their products reach customers on time.	https://www.kaggle.com/datasets/prachi13/customer-analytics/data	CSV	440.46 KB	Public