

Data Collection and Preprocessing Phase

Date	8 JULY 2024
Team ID	SWTID1720108776
Project Title	Ecommerce 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	The project estimates that, whether the product reached on time, which is based on various factors such as warehouse, the shipping method, the carrier used for shipping, and any potential delays or issues that may arise during the shipping process. This is Machine Learning Classification approach.
Data Collection Plan	From Kaggle- “Ecommerce Shipping data” the dataset is going to be collected.
Raw Data Sources Identified	<p>Name of the Dataset in Kaggle: E-Commerce Shipping Data, File Size: 124KB</p> <p>The collected data contains variables such as</p>

	Mode_of_Shipment, Cost_of_the_Product, Product_importance, Weight_in_gms, Reached.on.Time_Y.N etc
--	----------------------------------------------------------------------------------------------------------

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	The dataset includes shipment details like Mode_of_Shipment, Cost_of_the_Product, Product_importance, Weight_in_gms, Reached.on.Time_Y. N etc	https://www.kaggle.com/datasets/prachi13/customer-analytics?select=Train.csv	CSV	124 KB	Public