

Data Collection and Preprocessing Phase

Date	7 July 2024
Team ID	SWTID1720161281
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	Predicting whether, the product was shipped on time or not
Data Collection Plan	The data will be collected through Kaggle source using API keys
Raw Data Sources Identified	The source data consist of the 10999 entries of a product with product id an 11 other columns

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle	The data consist of 10999 rows and 12 columns having 12 separate variables	https://www.kaggle.com/datasets/prachi13/customer-analytics?select=Train.csv	CSV	440.46 kB	Public