

Data Collection and Preprocessing Phase

Date	4 JUNE 2024
Team ID	SWTID1720183095
Project Title	Ecommerce Shipping Prediction Using Machine Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

A few crucial actions must be taken in order to develop a plan for gathering data and locate sources of raw data for an e-commerce shipping prediction model. A thorough template to help you with the procedure is provided below:

Data Collection Plan Template

Section	Description
Project Overview	The goal of this project is to create a predictive model that will enable precise shipping time estimations for online orders. Through the use of machine learning algorithms, real-time data, and historical data, the model will increase the effectiveness of logistics while offering clients accurate delivery estimates.
Data Collection Plan	The dataset includes information collected from Kaggle.

Raw Data Sources Identified	We estimated times and anticipated on-time delivery using order data (items, locations, and times) from Kaggle.
-----------------------------	---

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	440KB	Public