

Data Collection and Preprocessing Phase

Date	25 June 2024
Team ID	SWTID1720434734
Project Title	Ecommerce shipping prediction
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 focuses on predicting shipping times for e-commerce orders by analyzing customer and order data. The goal is to enhance delivery efficiency and customer satisfaction by leveraging predictive analytics
Data Collection Plan	Data will be collected from various e-commerce platforms, including order details, customer demographics, and shipping information. Additional data sources will include historical delivery times, and third-party logistics information to ensure comprehensive analysis.
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle, the popular platform for data science competitions and repositories. The provided sample data represents a subset of the collected information ,encompassing variables related to customers

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle dataset	The dataset comprises of details about the customers and the time the product is shipped to the destination, cost of the products.	https://www.kaggle.com/datasets/prachi13/customer-analytics	CSV	431 KB	Public