Data Collection and Preprocessing Phase

Date	**10 July 2024**
Team ID	**SWTID1720163281**
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 "Ecommerce Shipping Prediction Using Machine Learning" is developed to estimate the shipping time of any product in an e-commerce environment. Customer satisfaction and loyalty should be fetched through correct predictions of the time of delivery; this shall be realized through incorporation of machine learning techniques incorporated in the analysis of several factors that determine the time—product type, customer location, and methods of shipments. It not only tries to satisfy every customer but also works towards enriching their experience with the ecommerce platform by making sure that its products are delivered on time. This in turn acts as a guarantee for the success and growth of the ecommerce business.

Data Collection Plan	Sales data: Extract all data that will include product ID, order date, shipping date, delivery date, customer location, warehouse location, and shipping method from the sales records of the company. Product Info: This includes the weight of the sold products, dimension, category
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle the popular platforms for data science competitions and repositories. The provided sample data represents a subset of the collected information, ware house block, customer care calls, cost of product , customer rating, prior purchase, product importance, reached on time, discount offer , Gender , mode of shipment machine learning analysis.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	ware house block, customer care calls, cost of product , customer rating, prior purchase, product importance, reached on time, discount offer , Gender , mode of shipment	https://www.kaggle.com/datasets/prachi13/customer-analytics?select=Train.csv	CSV	440 KB	Public