

Data Collection and Preprocessing Phase

Date	10 July 2024
Team ID	739758
Project Title	Walmart Sales Analysis For Retail Industry With Machine Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Report:

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 endeavour.

Data Collection Plan:

Section	Description
Project Overview	The machine learning project aims to predict sales for Walmart stores based on historical sales data and various influencing factors. Using a dataset with features such as store , department, holiday, weekend days, promotional events, temperature, unemployment, Weekly_Sales and historical sales; the objective is to build a model that accurately forecasts sales, facilitating efficient and informed decision-making in inventory management, staffing, and marketing strategies.
Data Collection Plan	Search for datasets related to retail sales, holidays, and promotional events. Prioritize datasets with detailed sales information and diverse influencing factors.

Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Walmart's internal sales records, Kaggle, and other relevant data repositories. The provided sample data represents a subset of the collected information, encompassing variables such as store location, sales figures, holiday periods, and promotional events for machine learning analysis.
-----------------------------	--



Raw Data Sources Report:

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	1.store dataset contains store ,type ,size 2.train dataset contains store,Dept,Date,Weekly_Sales,IsHoliday 3.feature dataset contains store,date,temperature,Fuel_Price,MarkDown1,MarkDown2,MarkDown3,MarkDown4,MarkDown5,CPI,unemployment,IsHoliday	https://www.kaggle.com/competitions/walmartrecruitingstore-salesforecasting/data	csv	3.2 MB	Public