Data Collection and Preprocessing Phase

Date	12 July 2024
Team ID	SWTID1720527361
Project Title	Traffictelligence-Advanced-Traffic-Volume Estimation-With-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	This machine learning project aims to predict traffic volume using weather, holiday, and time data. By analyzing historical data, the project will provide insights for better traffic management and urban planning. The ultimate goal is to develop a high-performing model that accurately forecasts traffic patterns, aiding in the optimization of traffic flow and reduction of congestion.

Data Collection Plan	Sources: Kaggle • Search for Datasets: • Go to Kaggle and use search terms like "traffic volume demographics" or "traffic prediction dataset." • Notable Datasets:
	• Metro Interstate Traffic Volume: While it primarily focuses on traffic data, combining it with external demographic data (e.g., from census data) can enrich the analysis. • Traffic Volume Counts: This dataset may include demographic factors influencing traffic patterns. Skill Wallet Enrolled for the project Trafficintelligence under Guided projects after that you can access datassets. under Data processing under skill wallet Utilize Skill Wallet to track progress in learning these skills through the courses and projects.
Raw Data Sources Identified	**Kaggle:** Traffic volume data with temperature, weather, holidays, and timestamps. Traffic counts at intersections, possibly with demographic data. SkillWallet Learn to integrate traffic data with demographics and apply skills to projects Track progress in data preprocessing, merging, and feature engineering

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions

| Kaggle | • **Traffic Volume**
 • Metro Interstate Traffic Volume: Number of vehicles over time.
 • **Weather Conditions** • Temperature , precipitation , wind speed, cloud cover, weather descriptions.
 • **Temporal Features**
 • Date and time stamps for extracting features like day of the week and hour. | Metro Interstate Traffic Volume | CSV | 0.0006 48498 5 GB | Public |

SkillWallet	Skill Wallet focuses on skill tracking and learning resources rather than hosting datasets. It offers skill	Dataset Link	Excel	0.0022 GB	Private (Create account and payment)
	development courses related to traffic analysis, covering topics like data preprocessing, dataset merging, and feature engineering. You can track your progress and apply learned skills to realworld projects involving traffic data integration and analysis..				