Data Collection and Preprocessing Phase

Date	4 July 2024
Team ID	SWTID1720090524
Project Title	Garment Worker Productivity 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 machine learning project aims to predict garment worker productivity based on user information. Using a dataset with features such as quarter, department, day, team number, time allocated, unfinished items, over time, incentive, idle time, idle men, style change, number of workers, and actual productivity, the objective is to build a model that accurately predicts productivity, facilitating efficient and informed decision-making in the process.
Data Collection Plan	● Search for datasets related to garment worker productivity. ● Prioritize datasets with diverse information.
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle & UCI, the popular platforms for data science competitions and repositories. The provided sample data represents a subset of the collected

	information, encompassing variables such as quarter, department, day, team number, time allocated, unfinished items, over time, incentive, idle time, idle men, style change, number of workers, and actual productivity for machine learning analysis.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	Our dataset contains information on various attributes of garment production, including the quarter, department, day, team number, time allocated, unfinished items, over time, incentive, idle time, idle men, style change, number of workers, and actual productivity.	https://www.kaggle.com/datasets/ishadss/productivity-prediction-of-garment-employees	CSV	93 KB	Public