

Data Collection and Preprocessing Phase

Date	19 July 2025
ProjectName	Machine Learning Approach for Employee Performance Prediction
Maximum Marks	

Data Quality Report:

The Data Quality Report will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Quality Report:

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	null values in 'wip'	Low	Removed the column
Kaggle Dataset	Two similar categories in 'department'	Moderate	merged into one category
Kaggle Dataset	Categorical data in the dataset	Moderate	Encoding has to be done in the data