Data Collection and Preprocessing Phase

Date	12 February 2026
Team ID	LTVIP2026TMIDS50820
Project Title	Prosperity Prognosticator: Machine Learning for Startup Success Prediction
Maximum Marks	2 Marks

Data Quality Report

The Data Quality Report Template 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 Source: Startup companies

Data Source	Data Quality Issue	Severity	Resolution Plan
Status column	Target column is in Categorical value	High	Converted into integer type as: Acquired = 1 Closed =0
age_first_milestone_year column	Missing values in the age_first_milestone_year filed	Moderate	Filling or Replacing missing values with 0

age_last_milestone_year column	Missing values in the age_last_milestone_year field	Moderate	Filling or Replacing missing values with 0.