

Data Collection and Preprocessing Phase

Date	17 July 2025
Team ID	739816
Project Title	Galactic Gazetteer: A comprehensive Dataset of Planet images
Maximum Marks	2 Marks

Data Quality Report Template

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	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	Missing/Corrupted Images	High	Check for corrupted images using verification methods; re-download or remove them.
Kaggle Dataset	Inconsistent Image Sizes	High	Resize all images to a consistent size (e.g., 224x224) or apply padding/cropping.
Kaggle Dataset	Incorrect Labeling	High	Verify labels manually or through metadata; correct or remove erroneous labels.