

Data Collection and Preprocessing Phase

Date	21 June 2024
Team ID	739940
Project Title	Gem Valuation Revolution: Predicting Diamond Prices With Artificial Neural Networks
Maximum Marks	2 Marks

Data Quality Report:

By systematically assessing data quality and preparing the dataset appropriately, you ensure that the predictive model for diamond prices is built on reliable, accurate data, leading to more trustworthy insights and decisions within the diamond industry.

Data Quality Report:

Data Source	Data Quality Issue	Severity	Resolution Plan
Diamond Dataset	Outliers in 'price' and 'carat' attributes detected beyond statistical thresholds.	High	Apply robust outlier detection methods (e.g., IQR, Z-score) and consider domain knowledge to determine validity or remove outliers.
Diamond Dataset	Potential data duplication or inconsistency across records.	Low	Conduct data deduplication and ensure data integrity through unique identifier checks (e.g., diamond ID).