

Ans: @ why do we preprocess the data:

Data preprocessing is a crucial step in the data mining and machine learning algorithms pipelines. It involves cleaning, transformation, and organising raw data into a format suitable for analysis. Several reasons highlight the importance of data preprocessing.

1. Handling missing values.
2. Dealing with Noisy Data
3. Data integration
4. Data transformation
5. Handling inconsistencies
6. Dimensionality Reduction
7. Addressing Skewness & Non-Normality
8. Improving Model performance
9. Enhancing interpretability
10. Meeting Algorithm Requirements.

→ Different algorithms have specific data requirements. Data preprocessing ensures that the input data meets these requirements, leading to optimal algorithm performance.