

# Data Readiness Checklist

The four problems that kill AI projects

---

Before feeding data to AI, check for these common issues. Poor data quality matters more than which model you choose.



Data follows different formats, naming conventions, or structures across sources. AI gets confused by conflicting patterns.



Missing values, gaps in time series, or incomplete records. AI can't work with what isn't there.



Data is outdated or no longer reflects current reality. Decisions based on old data lead to wrong conclusions.



Important metadata, relationships, or business logic is lost. AI sees isolated facts without understanding their meaning.

- Standardize formats  
Dates, currencies, and units are consistent throughout
- Handle missing values  
Decide: impute, remove, or flag missing data appropriately
- Remove duplicates  
Identify and handle duplicate records without losing information
- Validate freshness  
Ensure data reflects current state, especially for time-sensitive decisions
- Preserve context  
Include relevant metadata, relationships, and business rules
- Sample test  
Try AI on a small sample before processing the full dataset

 **The 80/20 Rule**

**80% of AI project time should be spent on data preparation.**

If you're spending more time choosing models than cleaning data, you're probably doing it wrong. Start with a small, clean dataset and expand once you prove the concept works.