



## **Handling Missing Values:**

Colab:

<https://colab.research.google.com/drive/1YV33jUv9qG5dAN3TNRs7ANCruR29-pZf>

### **What Are Missing Values?**

Missing values occur when no data is stored for a variable in an observation (e.g., NaN, NULL, ?, empty).

### **Types of Missing Data**

1. MCAR (Missing Completely at Random)
  - Missingness unrelated to any feature or target
  - Least problematic
2. MAR (Missing at Random)
  - Missingness related to other observed variables
  - Common in real-world datasets
3. MNAR (Missing Not at Random)
  - Missingness related to the missing value itself
  - Hardest to handle (requires domain knowledge)

### **Why Missing Values Are a Problem**

- Many ML algorithms cannot handle missing values directly
- Bias in predictions
- Reduced statistical power
- Can distort distributions