



# Data Cleaning

Data cleaning, also known as data cleansing or data scrubbing, is a crucial step in the data analysis process. It involves identifying and correcting errors, inconsistencies, and inaccuracies in datasets to enhance their quality and reliability.

# Introduction

01

Errors and inconsistencies can be corrected to improve dataset quality

02

Data cleaning is essential for accurate and reliable analysis

03

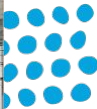
It ensures that information reflects the real world accurately





# Accuracy

- Cleaning data ensures the information it contains is correct
- Like proofreading an essay to communicate the right message
- Ensures all puzzle pieces fit together to reveal the whole picture



# Consistency

01

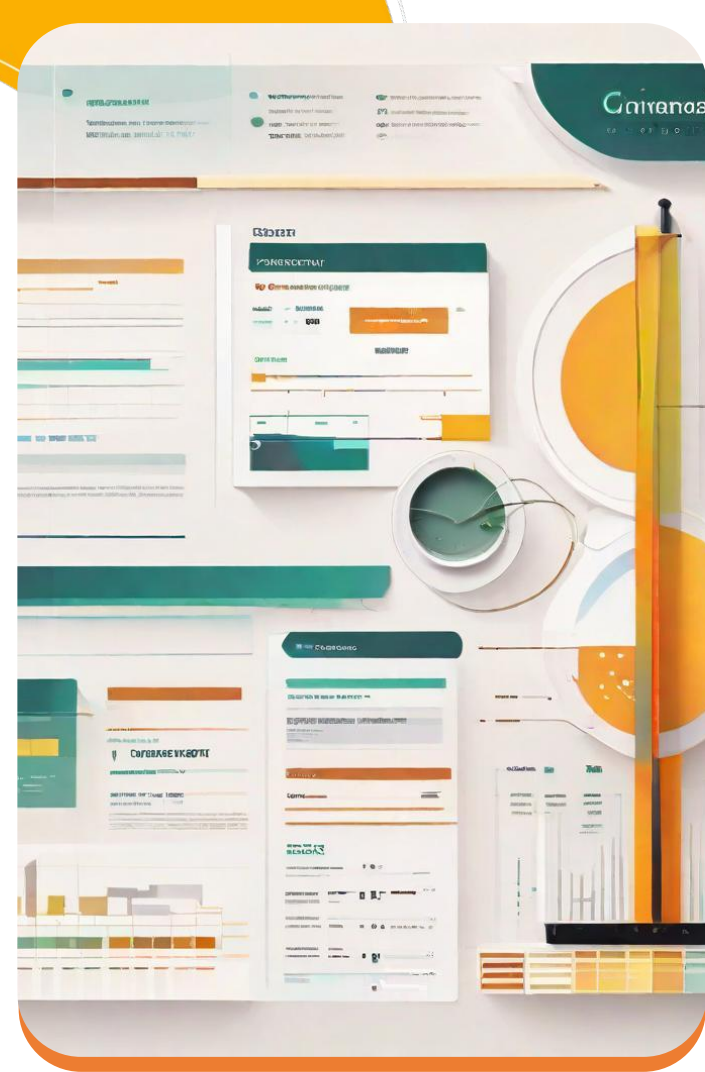
Like using the same measurements in recipes for clarity

02

Standardizing formats and resolving differences

03

Ensures all information is uniform and easily compared or combined



# Completeness

01

Similar to having all the ingredients to cook a delicious meal

02

Addressing missing values and outliers

03

Ensures we have a complete dataset for clearer understanding





# Reliability

01

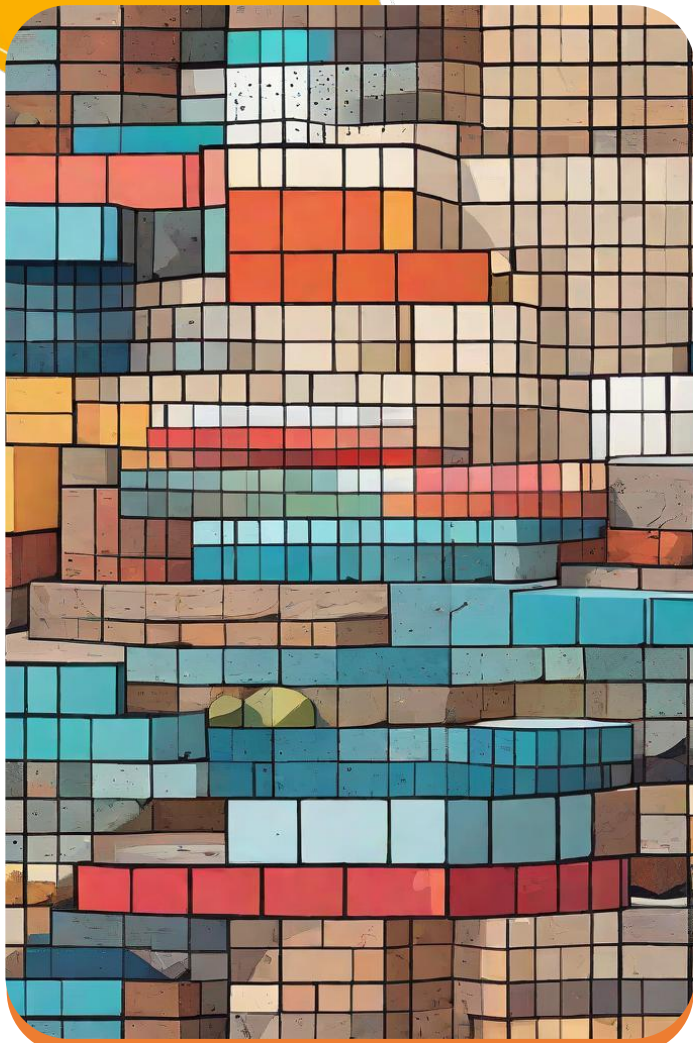
Similar to relying on a reliable friend for guidance

02

Clean data helps make informed decisions and draw meaningful conclusions

03

Ensures information is accurate and trustworthy





# Data Integrity

- Identifying and fixing errors to maintain trustworthiness of analysis
- Ensures insights are based on solid, dependable data
- Similar to a security guard checking IDs at a party



# Conclusion

- 01 Never underestimate the importance of data cleaning
- 02 Data cleaning is the unsung hero of the data analysis process
- 03 Ensures insights drawn from datasets are accurate, reliable, and actionable





**Thank you for your time 😊**