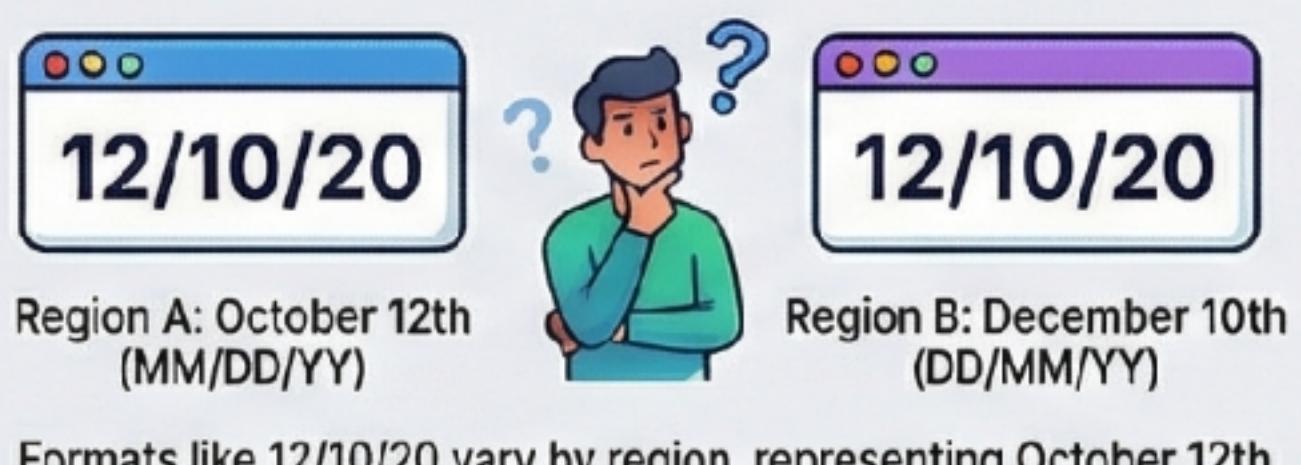


The Foundation of Trust: Navigating Global Data Integrity

The Standardization Challenge

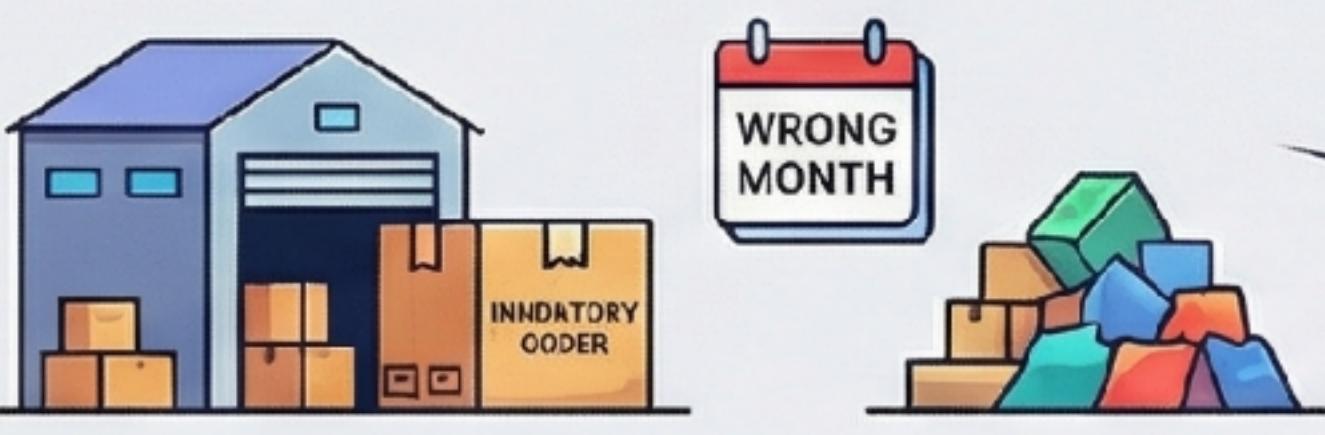
Global Date Ambiguity



Formats like 12/10/20 vary by region, representing October 12th or December 10th depending on local standards.



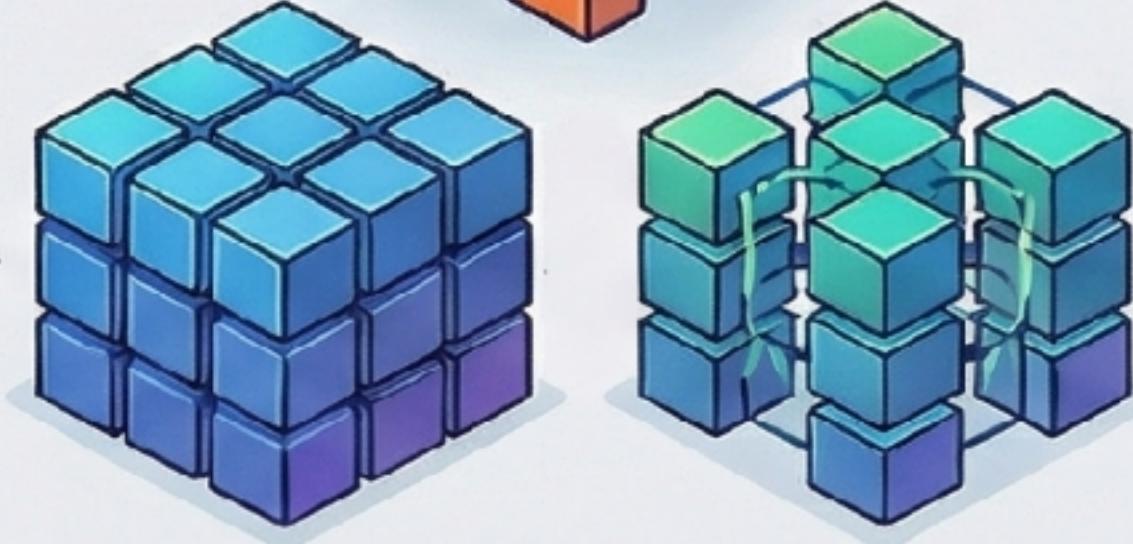
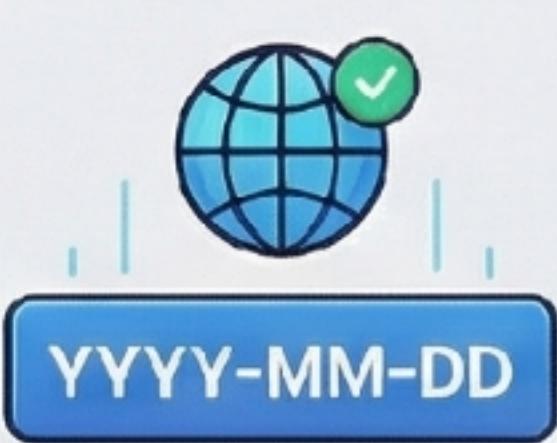
The Cost of Inconsistency



Failing to standardize formats can lead to critical errors, such as ordering inventory for the wrong month.

The Standard Solution

Adepting a common format like YYYY-MM-DD ensures data remains valid across international borders.

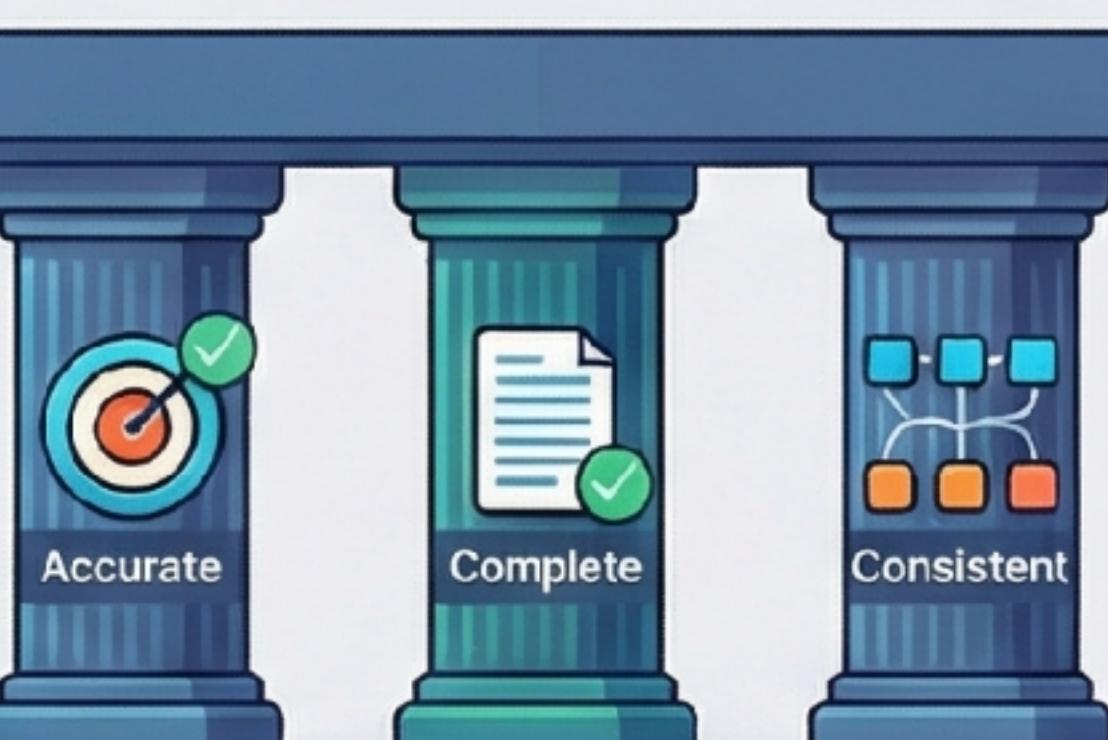


Safeguarding Data Quality



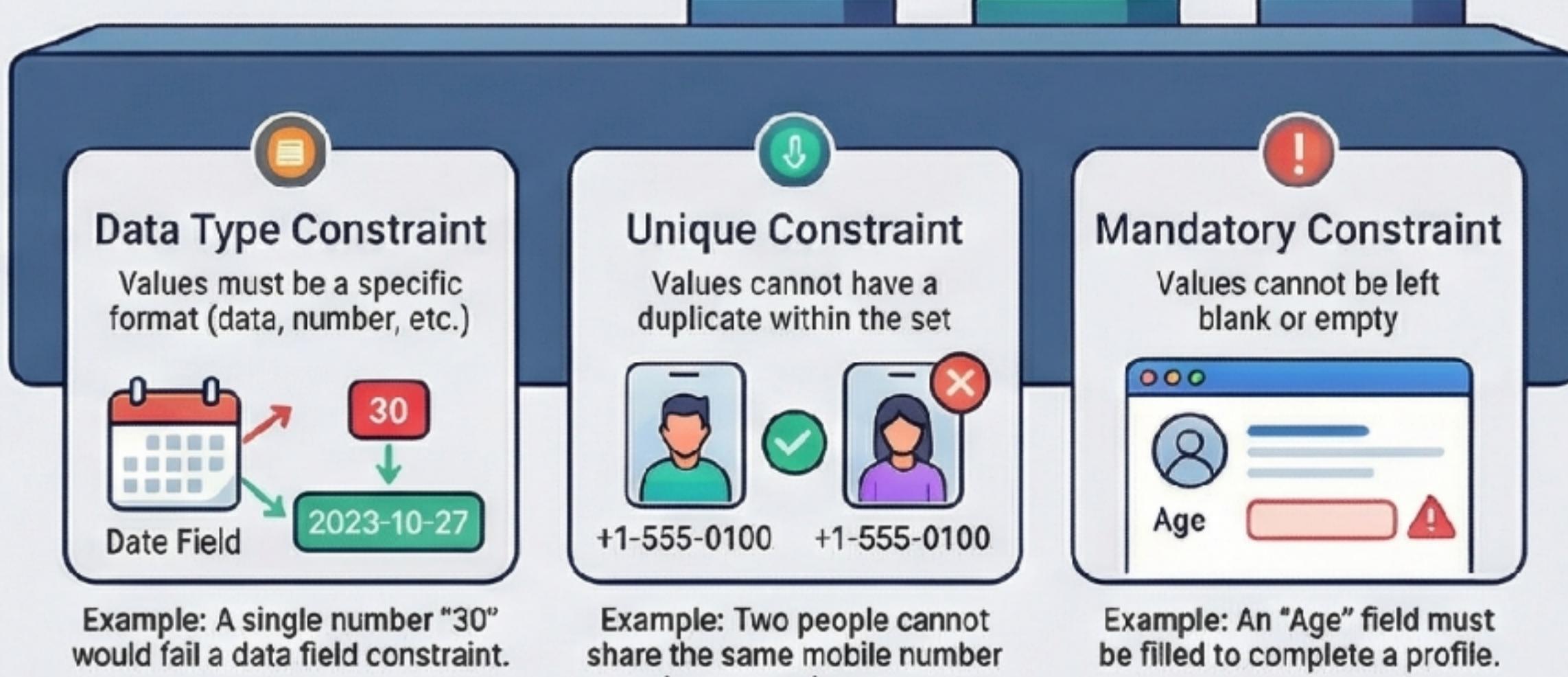
Integrity is often compromised during data replication, improper field classification during transfers, or accidental record deletion.

Mitigating Processing Risks



The Pillars of Data Health

High-quality data must be accurate, complete, and consistent across all entry points and databases.



Data Type Constraint

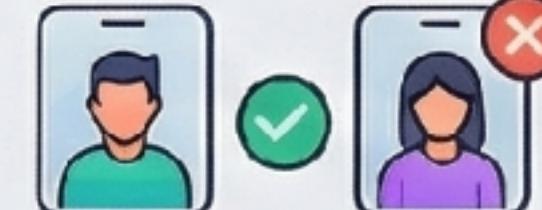
Values must be a specific format (data, number, etc.)



Example: A single number "30" would fail a data field constraint.

Unique Constraint

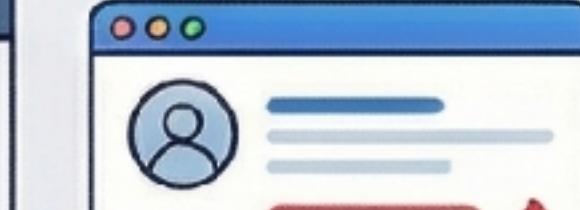
Values cannot have a duplicate within the set



Example: Two people cannot share the same mobile number in one service area.

Mandatory Constraint

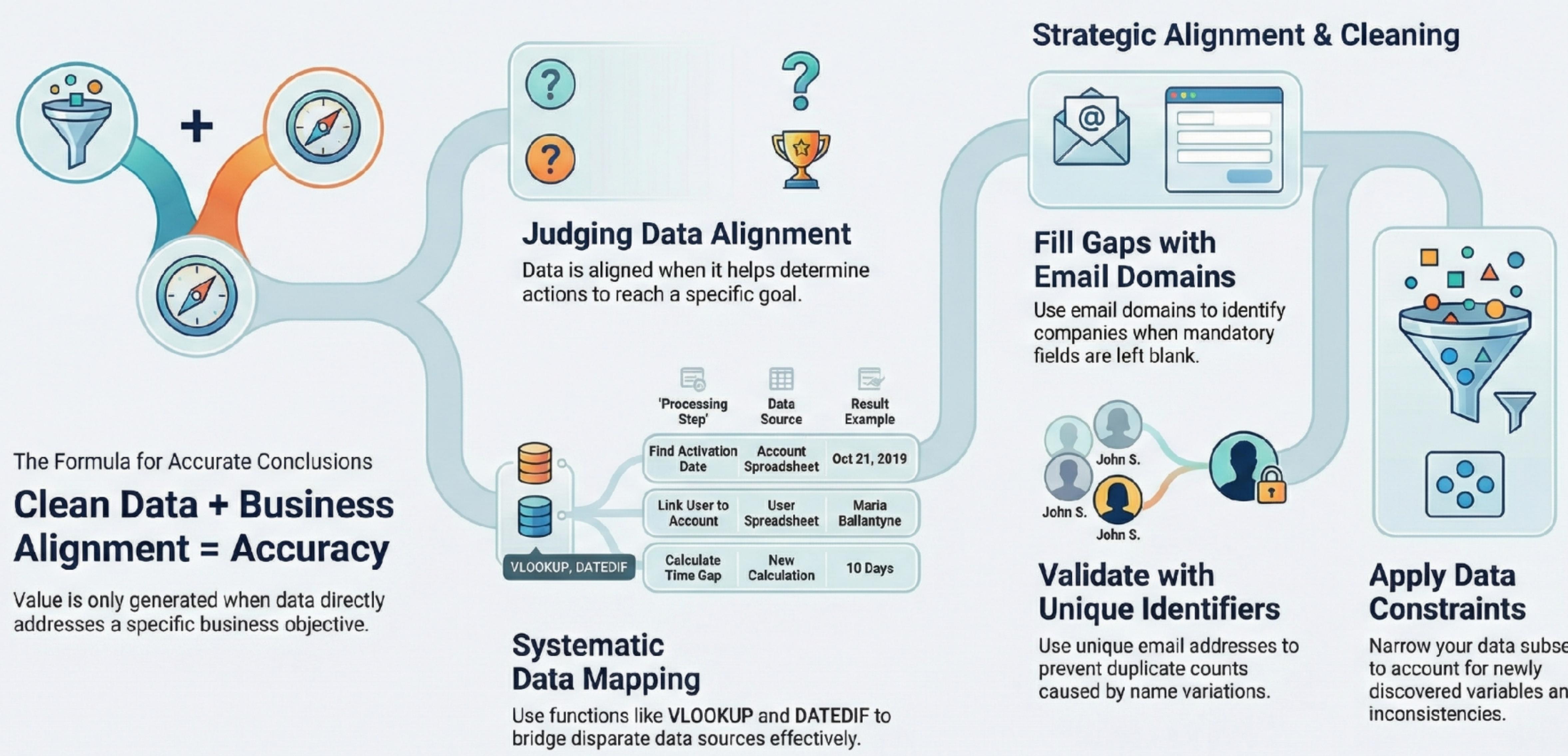
Values cannot be left blank or empty



Example: An "Age" field must be filled to complete a profile.



The Path to Precise Insights: Aligning Data with Business Goals



Data Roadblocks: Navigating Gaps and Errors

SOLVING MISSING OR INSUFFICIENT DATA

Utilize Proxy Data

Substitute missing information with similar datasets, such as data from a comparable city or demographic.



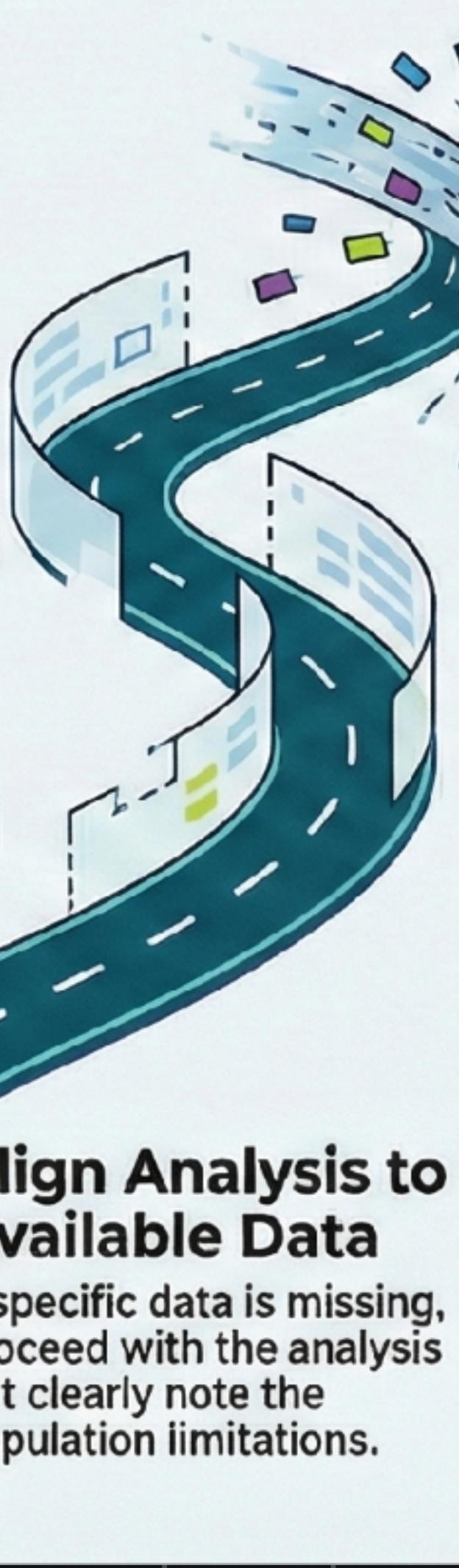
Small-Scale Collection

Conduct a preliminary analysis on a small sample while requesting more time for full collection.



Align Analysis to Available Data

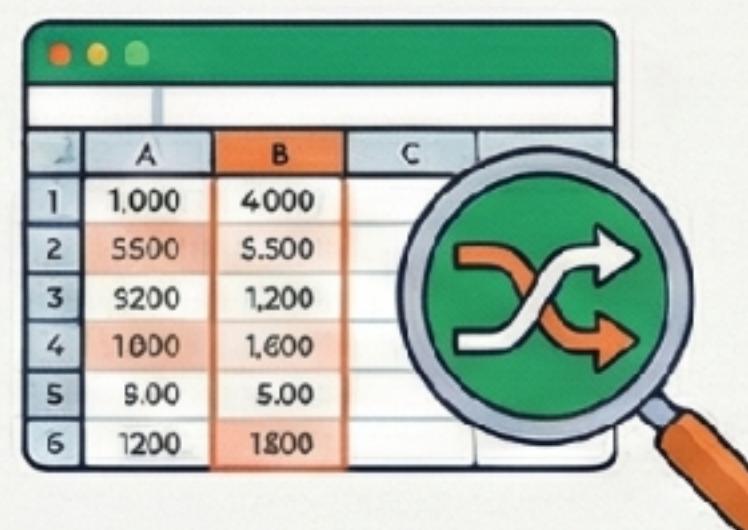
If specific data is missing, proceed with the analysis but clearly note the population limitations.



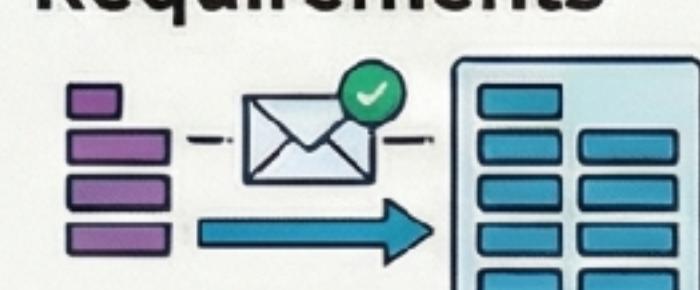
MANAGING ERRONEOUS DATA

Identify Pattern-Based Errors

Fix errors at the source by correcting underlying logic or conditional statements in spreadsheets.



Re-Communicate Requirements



If data was misinterpreted, restate your specific needs to the data provider to get a fresh set.

Strategic Omission

Ignore bad data only if the remaining sample is large enough and free of systematic bias.

