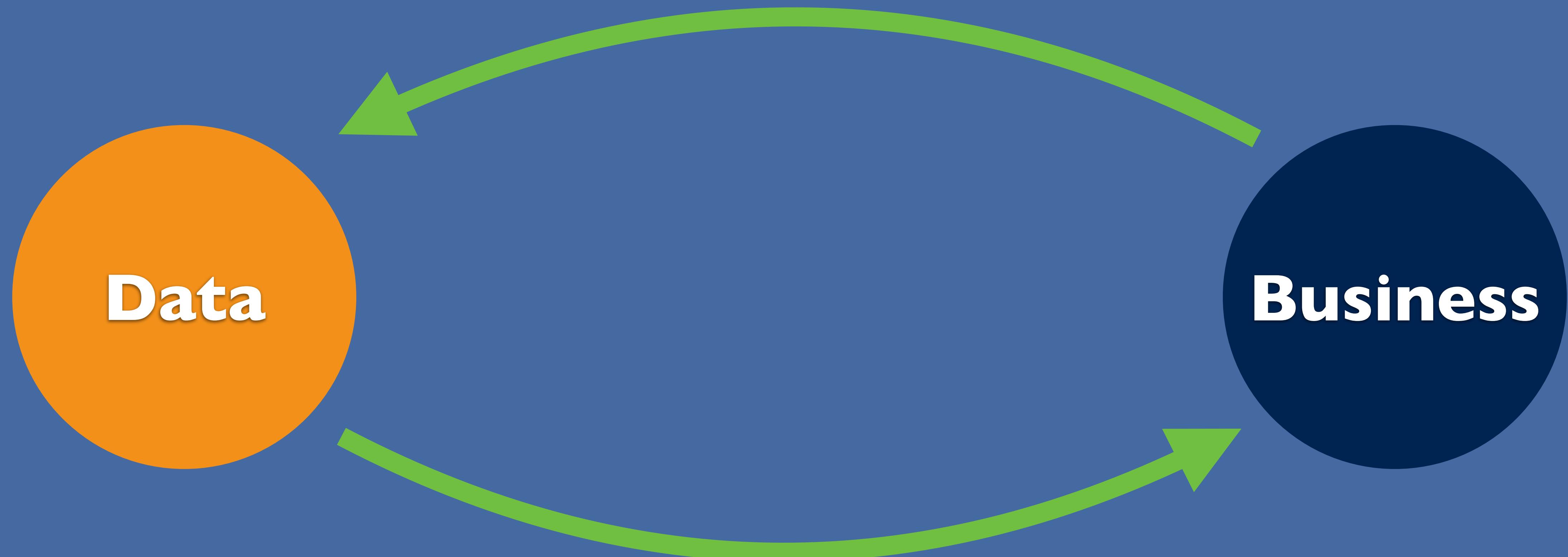


# DATA & BUSINESS

Business to Data = B2D

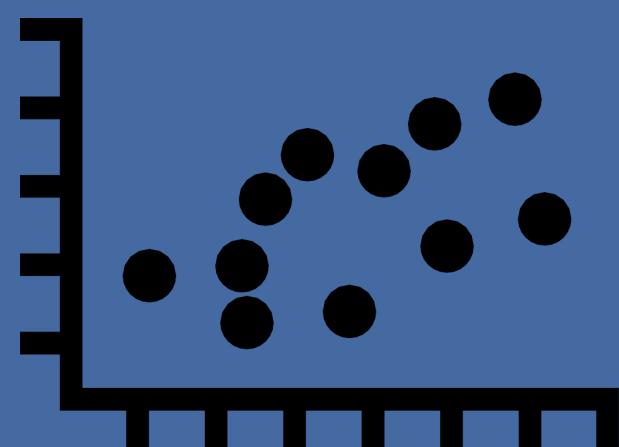


Data to Business = D2B

# FROM DATA TO BUSINESS

Example

Air  
Pressure



Data



Information

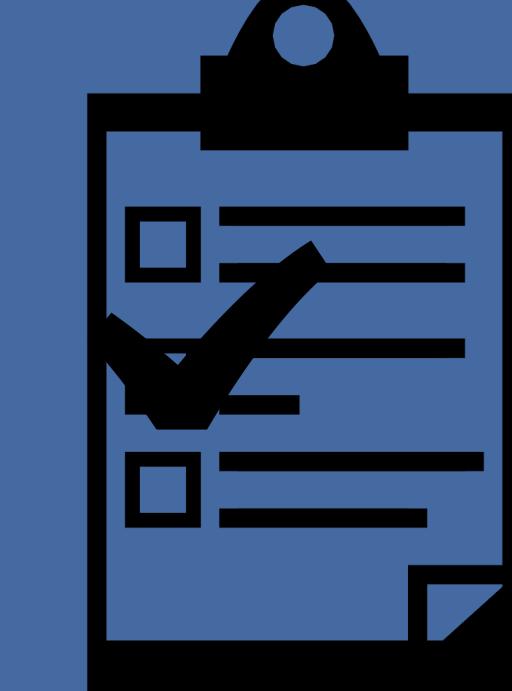
Evolution



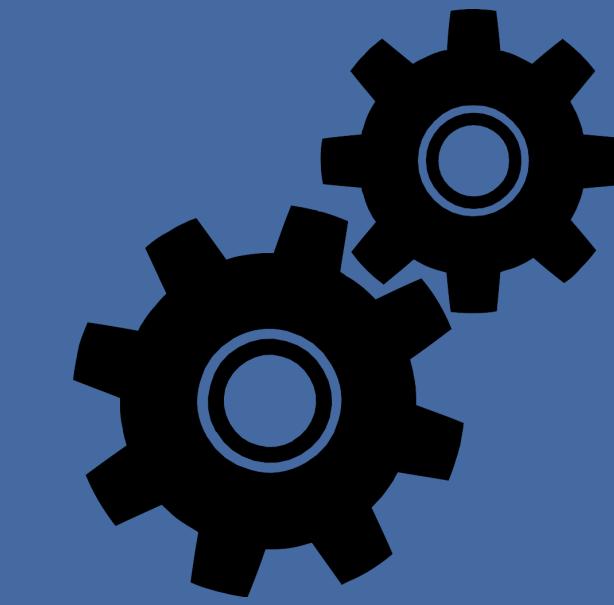
Analysis

Increase  
Decrease

Push  
category



Decisions



Actions

Promotion

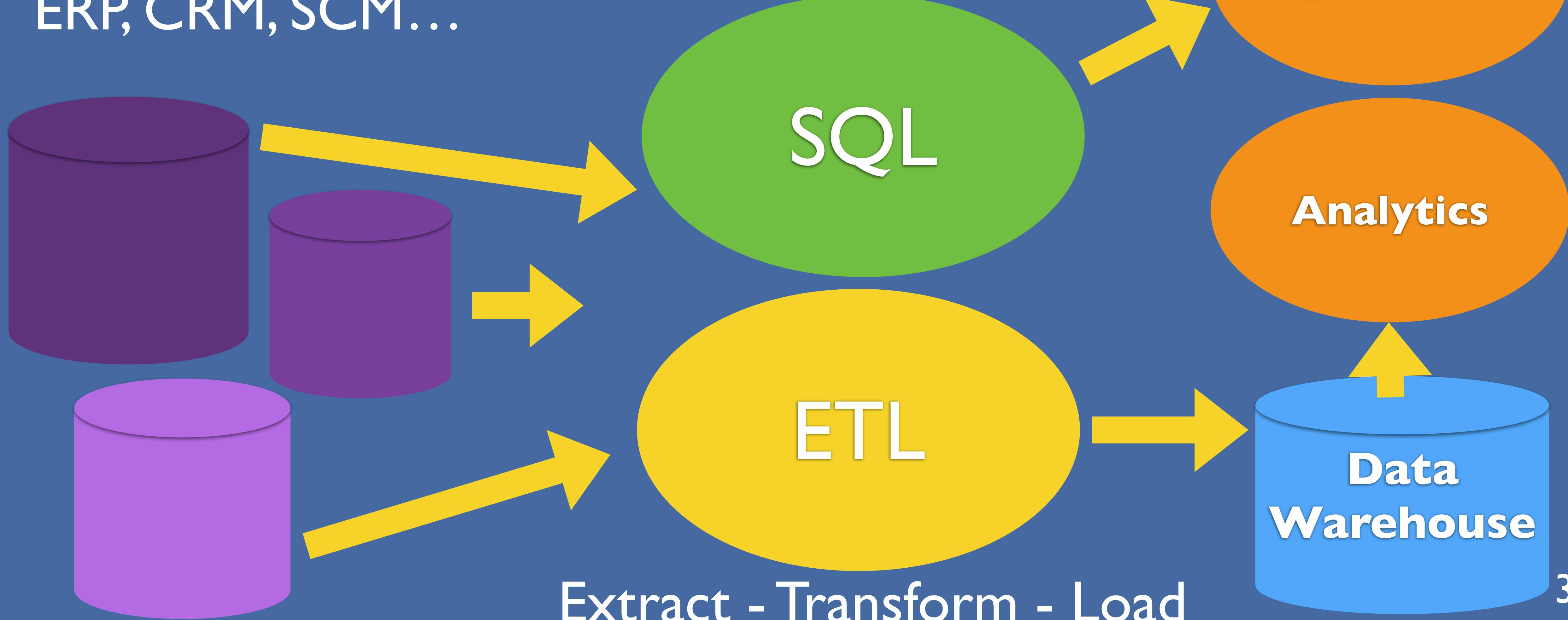


Business  
Value

Sales

# ACCESS TO DATA

Operation Databases:  
ERP, CRM, SCM...



# DATA WAREHOUSE & DATA LAKE

## Data Warehouse

Structured data

Transactions

Batch

Mostly internal datasets

## Data Lake

Structured + Unstructured data

Interactions

Real time

Internal & External datasets

# DATA PRE-PROCESSING

Quality

Compliance

Usability

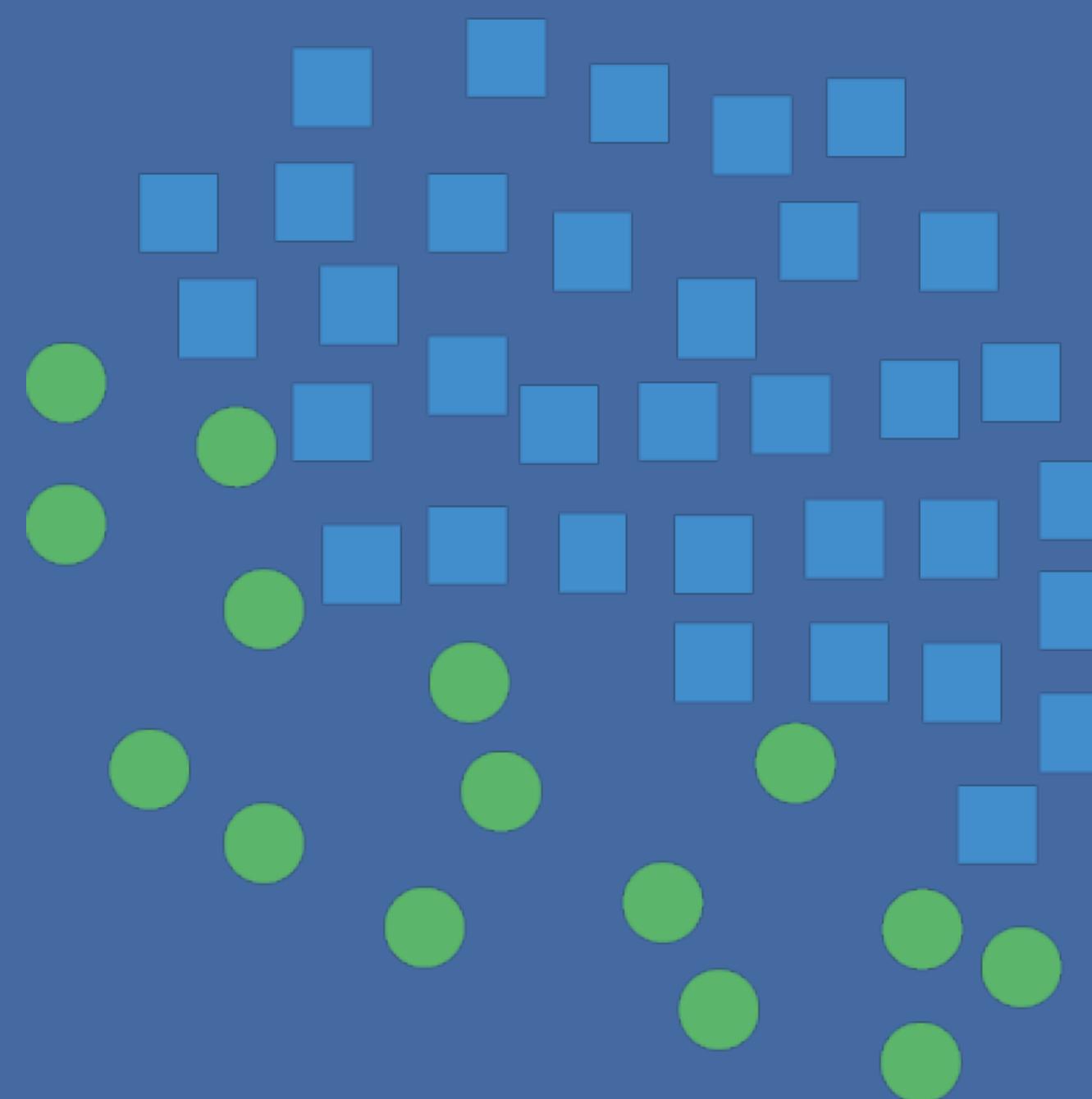
# DATA PRE-PROCESSING

Quality

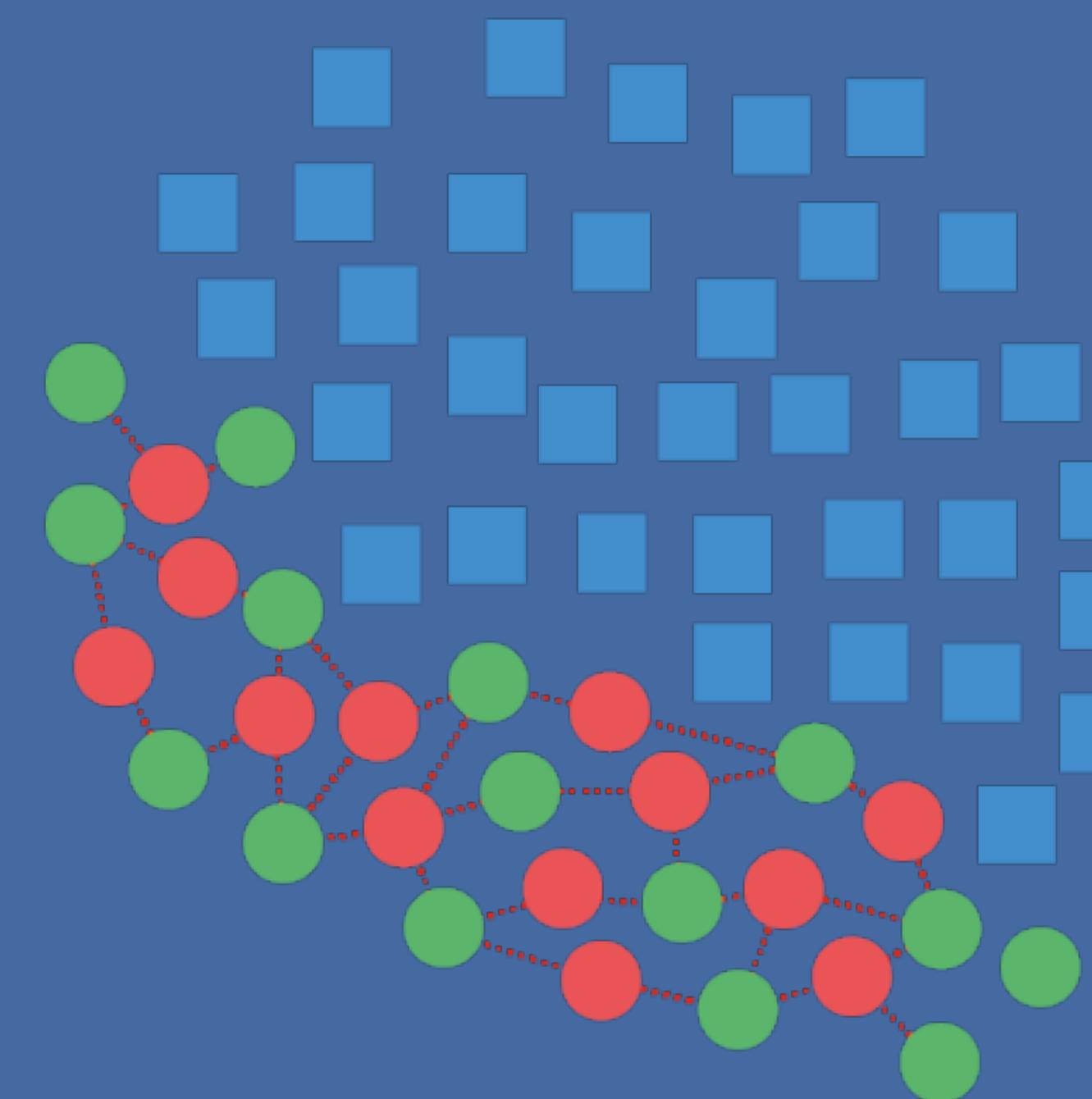
Missing values  
Errors in values  
Errors in formats  
Imbalanced data sets  
...

# OVERSAMPLING WITH SMOTE

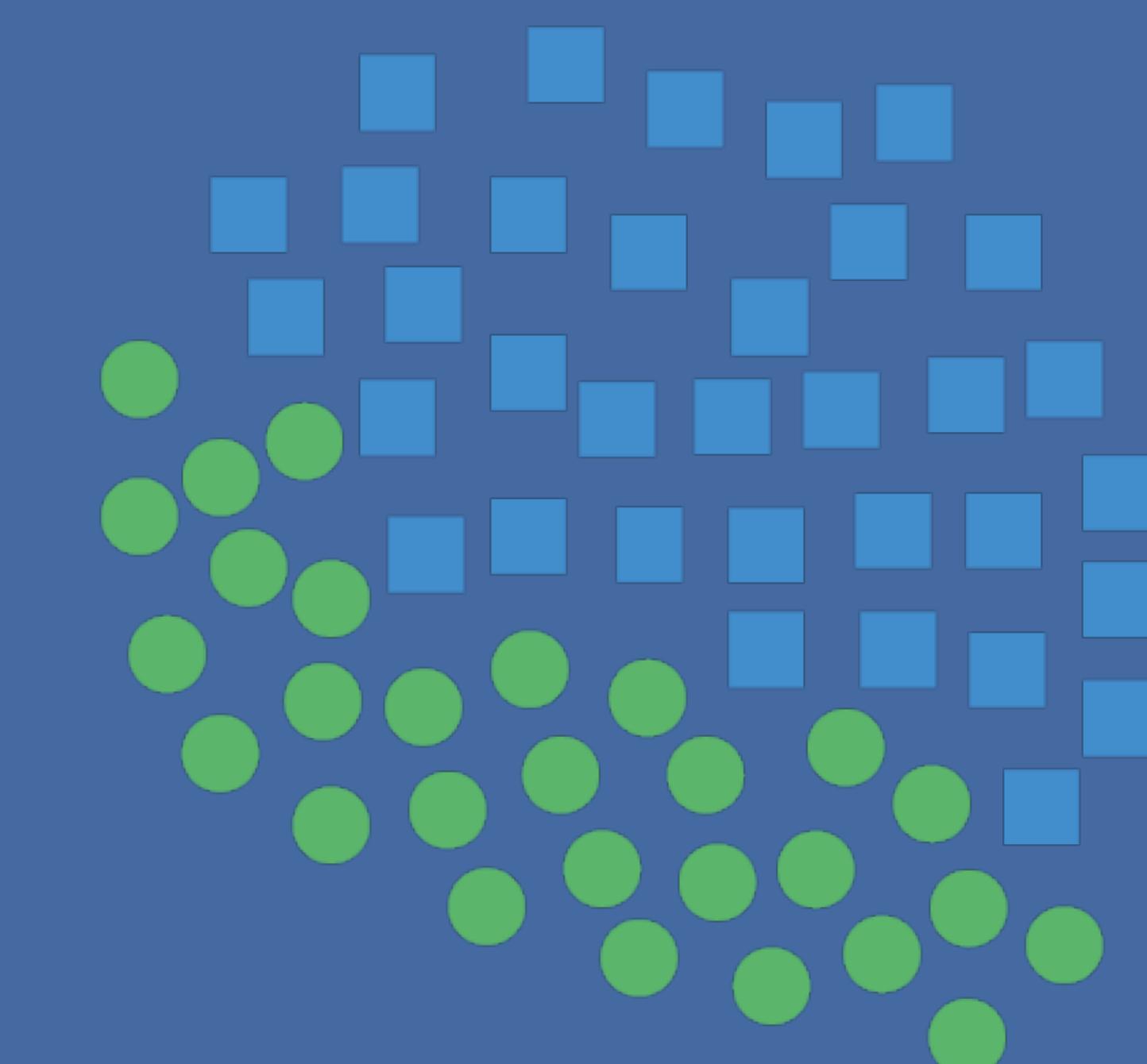
Synthetic Minority Oversampling Technique



Original Dataset



Generating Samples



Resampled Dataset

# DATA PRE-PROCESSING

Compliance

Legal constraints  
Sector constraints  
Internal constraints  
...

# DATA PRE-PROCESSING

Usability

Understandable for algorithms

Efficient for algorithms

Robust for algorithm

Automated for algorithm

...

→ Feature Engineering

# DATA ARCHITECTURE

