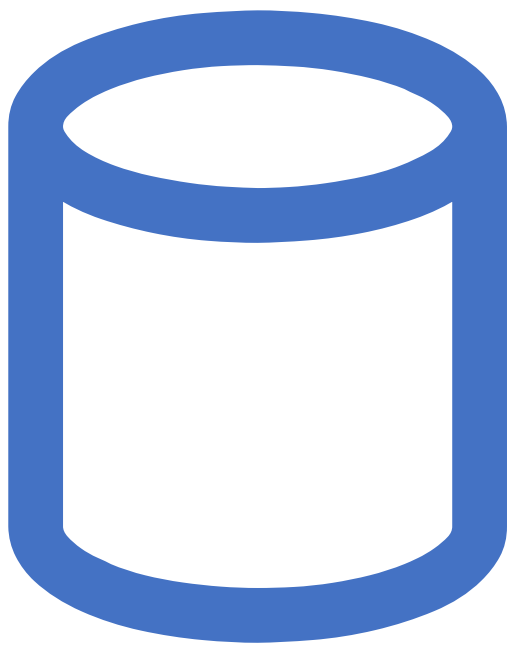
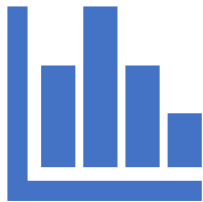




Praposes  
Data





Pengantar praposes  
data



Membersihkan data  
(data cleaning)



Mentranformasi data

# Kenapa perlu melakukan praposes

---



**Garbage In**



**Analysis**



**Garbage Out**

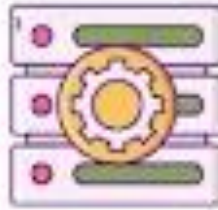
# Extract transform load

---



**Extract**

Mengumpulkan data dari berbagai sumber



**Transform**

Melakukan praproses data



**Load**

Menyimpan data

# Apa kebutuhan melakukan praproses data?

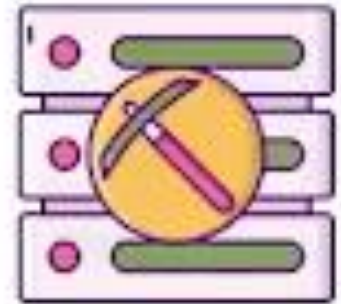
---



Penyimpanan data di  
*data warehouse*



Membuat laporan  
(*reporting*)



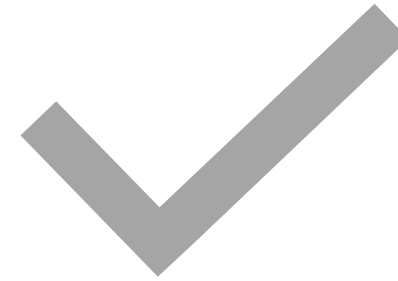
Analisis untuk  
menggali *insight*

# Tahapan melakukan praposes data

---

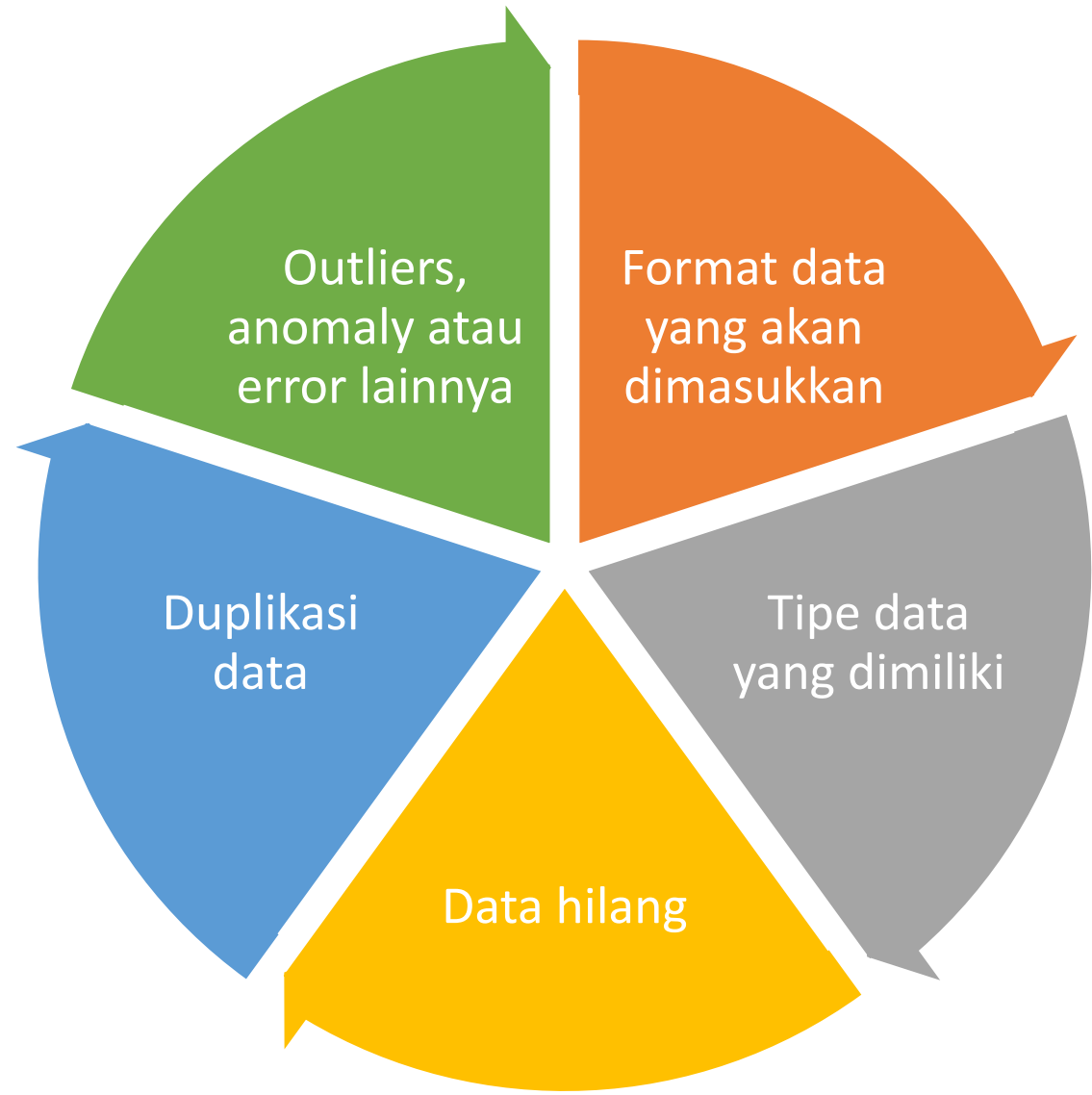


Membersihkan data



Transformasi data

# Pembersihan data



---

## Data input



**Excel**



**Comma  
Separated  
Value**



**Database**



**Javascript  
Object  
Nation**

# Praktik input data



# Tranformasi data

- Feature engineering

| Lender ID | Contract type | Amount    | Interest | Grade |
|-----------|---------------|-----------|----------|-------|
| 1         | C             | 1 000 000 | 20%      | A1    |
| 2         | C             | 2 000 000 | 20%      | A3    |
| 2         | S             | 5 000 000 | 10%      | B1    |
| 2         | S             | 3 000 000 | 10%      | B3    |
| 3         | C             | 1 000 000 | 20%      | C2    |
| 3         | S             | 2 000 000 | 10%      | A     |

| Profit  | Is conventional | Funding category | Normalize | New grade |
|---------|-----------------|------------------|-----------|-----------|
| 200 000 | 1               | Low              | 0         | A         |
| 400 000 | 0               | Normal           | 0.25      | B         |
| 500 000 | 0               | High             | 1         | B         |
| 300 000 | 0               | Normal           | 0.5       | C         |
| 200 000 | 1               | Low              | 0         | C         |
| 200 000 | 0               | Normal           | 0.25      | A         |

Extract

Encode

Discretize

Scale

Map

- Extract** : Membuat kolom baru dari operasi kolom lain
- Encode** : Mengubah kolom *string* menjadi angka
- Discretize** : Membuat kategori baru dari suatu kolom
- Scale** : Mengubah unit satuan dari data
- Map** : Mengubah angka yang diinginkan

# Praktik future engineering



# Mengubah bentuk tabel

---

| Funding ID | Lender ID | Contract type | Amount    |
|------------|-----------|---------------|-----------|
| 1          | 1         | C             | 1 000 000 |
| 2          | 2         | C             | 1 000 000 |
| 3          | 2         | S             | 1 000 000 |
| 4          | 2         | S             | 2 000 000 |
| 5          | 3         | C             | 1 000 000 |
| 6          | 3         | S             | 2 000 000 |

→  
Aggregating

| Lender ID | Total amount |
|-----------|--------------|
| 1         | 1 000 000    |
| 2         | 4 000 000    |
| 3         | 3 000 000    |

| Funding ID | Lender ID | Contract type | Amount    |
|------------|-----------|---------------|-----------|
| 1          | 1         | C             | 1 000 000 |
| 2          | 2         | C             | 1 000 000 |
| 3          | 2         | S             | 1 000 000 |
| 4          | 2         | S             | 2 000 000 |
| 5          | 3         | C             | 1 000 000 |
| 6          | 3         | S             | 2 000 000 |

→  
Pivoting

| Lender ID | C         | S         |
|-----------|-----------|-----------|
| 1         | 1 000 000 |           |
| 2         | 1 000 000 | 3 000 000 |
| 3         | 1 000 000 | 2 000 000 |

←  
Melting

# Praktik mengubah bentuk tabel

---

