

# **Laporan Final Project**

## **Business Intelligence**



**Disusun oleh:**

Romy Mahardika Pangestu Lazuardi (2341720129)

**TI\_3A**

**PROGRAM STUDI D4 TEKNIK INFORMATIKA  
JURUSAN TEKNOLOGI INFORMASI  
POLITEKNIK NEGERI MALANG  
2025**

# 1. Studi Kasus Pengelolaan Data (Data Penjualan – Northwind)

Studi kasus yang digunakan adalah **pengelolaan data penjualan pada perusahaan distribusi produk** menggunakan **Northwind Dataset** (data dummy publik).

<https://www.kaggle.com/datasets/emmanuelugbeh/northwind-orders-and-order-details>

**Sumber data dummy:**

- Kaggle (Northwind Dataset – open source & legal)
- Data tersedia dalam format CSV

**Tabel yang digunakan ( $\geq 4$  tabel):**

1. **customers**  
customerID, companyName, contactName, contactTitle, city, country
2. **products**  
productID, productName, quantityPerUnit, unitPrice, discontinued, categoryID
3. **orders**  
orderID, customerID, employeeID, orderDate, requiredDate, shippedDate, shipperID, freight
4. **order\_details**  
orderID, productID, unitPrice, quantity, discount
5. **employees**  
employeeID, employeeName, title, city, country, reportsTo
6. **categories**  
categoryID, categoryName, description
7. **shippers**  
shipperID, companyName

## 2. Studi Kasus Terpilih

**Studi Kasus:**

*Analisis Penjualan dan Kinerja Produk pada Perusahaan Distribusi*

**Alasan pemilihan:**

- Data lengkap (penjualan, pelanggan, produk, karyawan)
- Mudah diakses dan legal
- Cocok untuk perhitungan KPI (penjualan, produk terlaris, performa karyawan)
- Relasi data jelas untuk skema bintang

### 3. Pengambilan Data Dummy

#### Langkah:

1. Download dataset Northwind (CSV/SQL)
2. Identifikasi format data:
  - CSV: customers.csv, products.csv, orders.csv, order\_details.csv, dll
3. Validasi atribut:
  - Tidak kosong
  - Tipe data sesuai (tanggal, numerik, string)

### 4. Rancangan Skema Bintang (Star Schema)

#### Tabel Fakta

##### Fact\_Penjualan

- order\_id
- product\_id
- customer\_id
- employee\_id
- shipper\_id
- time\_id
- quantity
- unit\_price
- discount
- total\_penjualan

#### Tabel Dimensi

##### 1. Dim\_Pelanggan

- customer\_id
- company\_name
- contact\_name
- city
- country

##### 2. Dim\_Produk

- product\_id
- product\_name
- category\_name
- unit\_price
- discontinued

### 3. Dim\_Karyawan

- employee\_id
- employee\_name
- title
- city
- country

### 4. Dim\_Pengirim

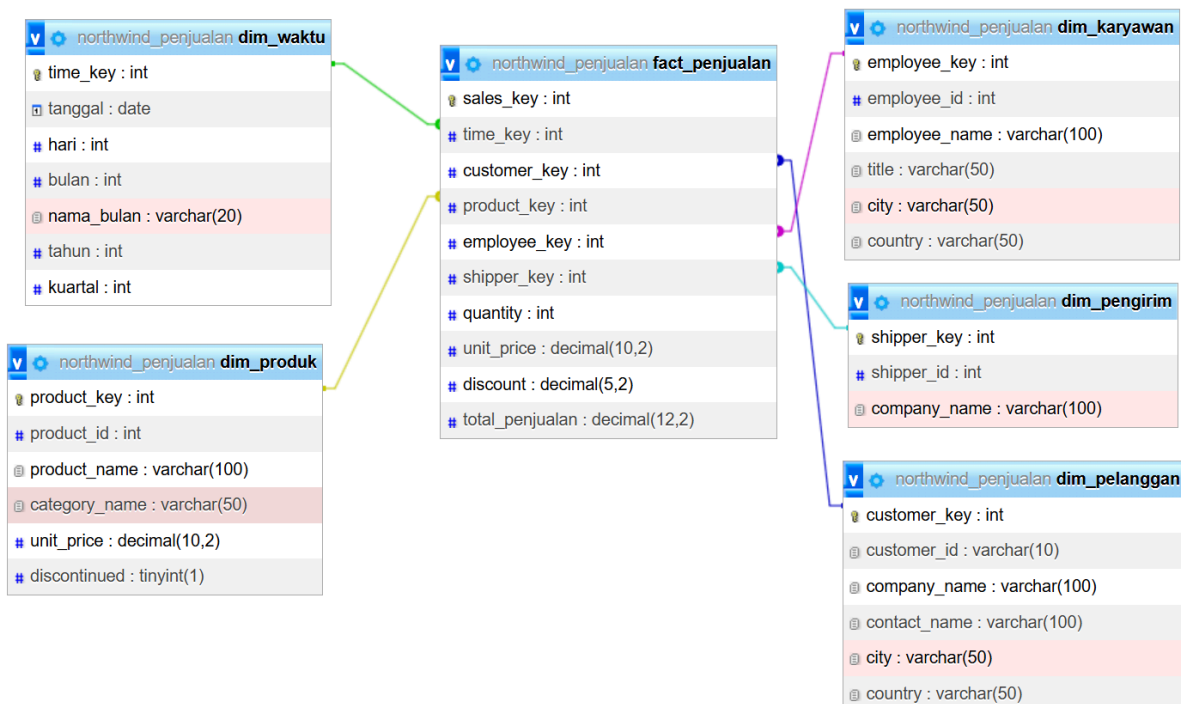
- shipper\_id
- company\_name

### 5. Dim\_Waktu

- time\_id
- tanggal
- bulan
- tahun
- kuartal

## Relasi

- Fact\_Penjualan memiliki **foreign key** ke seluruh tabel dimensi
- Relasi berbentuk **one-to-many**



# Rancangan Proyek ETL

## Extract

Data diekstrak dari file CSV Northwind menggunakan Pentaho dan dimuat ke staging area.

- Mengambil data dari:
  - File CSV (customers, orders, products, dll)
- Tools:
  - Pentaho Data Integration (Kettle)

## Transform

Dilakukan proses:

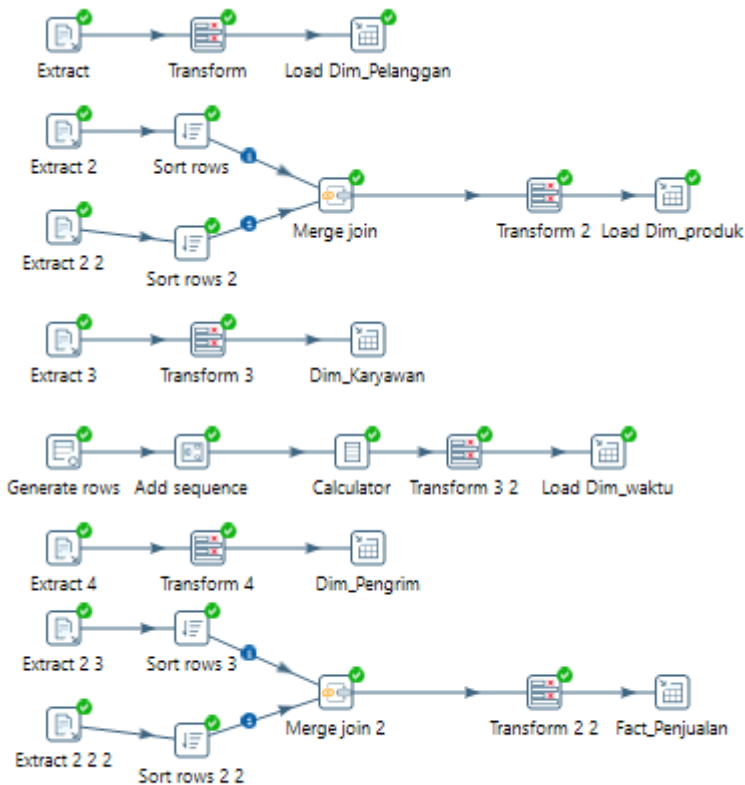
- Normalisasi data
  - Penyesuaian struktur agar sesuai skema bintang
  - Perhitungan nilai penjualan
  - Validasi data kosong atau tidak konsisten
- Cleaning data:
  - Menghapus duplikat
  - Standarisasi nama negara & kota
  - Konversi format tanggal
- Transformasi:
  - Hitung  $\text{total\_penjualan} = \text{quantity} \times \text{unit\_price} \times (1 - \text{discount})$
  - Join tabel orders + order\_details
  - Mapping kategori produk

## Load

Data yang telah bersih dimasukkan ke data warehouse:

- Dimensi → Fakta
- Relasi dijaga dengan foreign key
  - Load ke:
    - Tabel dimensi terlebih dahulu
    - Dilanjutkan tabel fakta

- Perhatikan:
  - Primary key
  - Foreign key
  - Integritas data



## Contoh Kasus

Total Penjualan

Showing rows 0 - 0 (1 total, Query took 0.0044 seconds.)

SELECT SUM(total\_penjualan) AS total\_penjualan FROM fact\_penjualan;

Profiling [ Edit inline ] [ Edit ] [ Explain SQL ] [ Create PHP code ] [ Refresh ]

Show all | Number of rows: 25 | Filter rows: Search this table

Extra options

total_penjualan
207306.10

Total Penjualan per Tahun

Showing rows 0 - 0 (1 total, Query took 0.0030 seconds.)

```
SELECT w.tahun, SUM(f.total_penjualan) AS total_penjualan FROM fact_penjualan f JOIN dim_waktu w ON f.time_key = w.time_key GROUP BY w.tahun;
```

☐ Profiling [\[ Edit inline \]](#) [\[ Edit \]](#) [\[ Explain SQL \]](#) [\[ Create PHP code \]](#) [\[ Refresh \]](#)

☐ Show all | Number of rows: 25 | Filter rows:

Extra options

tahun	total_penjualan
2018	8365.20

## Top 5 Produk Terlaris

Showing rows 0 - 4 (5 total, Query took 0.0025 seconds.)

```
SELECT p.product_name, SUM(f.quantity) AS total_terjual FROM fact_penjualan f JOIN dim_produk p ON f.product_key = p.product_key GROUP BY p.product_name ORDER BY total_terjual DESC LIMIT 5;
```

☐ Profiling [\[ Edit inline \]](#) [\[ Edit \]](#) [\[ Explain SQL \]](#) [\[ Create PHP code \]](#) [\[ Refresh \]](#)

Extra options

product_name	total_terjual
Gula Malacca	20
Steeleye Stout	20
Chef Anton's Gumbo Mix	18
Louisiana Hot Spiced Okra	17
Aniseed Syrup	16

