

# **LAPORAN TUGAS AKHIR BUSINESS INTELLIGENCE**

## **Implementasi ETL dan Data Warehouse pada Dataset Northwind Traders**



Anggota Kelompok:

Adinda Mirza Devani	(2341720046)
Dimas Adi Bayu Samudra	(2341720169)
Dwi Ahmad Khairy	(2341720073)
Riovaldo Alfiyan Fahmi R	(2341720209)
Zilan Zalilan	(244107023003)

**PROGRAM STUDI D-IV TEKNIK INFORMATIKA**

**JURUSAN TEKNOLOGI INFORMASI**

**POLITEKNIK NEGERI MALANG**

**2025**

## DAFTAR ISI

DAFTAR ISI.....	2
Link Github.....	2
BAB 1.....	3
IDENTIFIKASI STUDI KASUS.....	3
1.1 Latar Belakang & Pemilihan Studi Kasus.....	3
1.2 Kriteria Data.....	3
1.3 Tujuan Proyek.....	3
BAB 2.....	4
PERANCANGAN DATA WAREHOUSE (STAR SCHEMA).....	4
2.1 Struktur Skema.....	4
BAB 3.....	5
IMPLEMENTASI ETL (EXTRACT, TRANSFORM, LOAD).....	5
3.1 Implementasi Dimensi Produk (DimProduct).....	5
3.2 Implementasi Dimensi Pelanggan (DimCustomer).....	6
3.3 Implementasi Dimensi Karyawan (DimEmployee).....	7
3.4 Implementasi Tabel Fakta (FactSales).....	9
BAB 4.....	12
ANALISIS DAN VISUALISASI DASHBOARD.....	12
BAB 5.....	15
KESIMPULAN.....	15

## Link Github

<https://github.com/suckgitariuses/FIVIZ-BI>

# BAB 1

## IDENTIFIKASI STUDI KASUS

### 1.1 Latar Belakang & Pemilihan Studi Kasus

Penelitian ini berfokus pada optimasi pengelolaan data penjualan menggunakan studi kasus **Northwind Traders**. Perusahaan ini dipilih karena memiliki ekosistem data yang kompleks—mencakup transaksi, produk, pelanggan, dan karyawan—yang tersimpan dalam format CSV terpisah. Integrasi data melalui sistem *Business Intelligence* dianggap krusial bagi perusahaan retail seperti Northwind Traders untuk mengubah data mentah menjadi wawasan strategis dalam pengambilan keputusan bisnis.

### 1.2 Kriteria Data

Dataset yang diolah memiliki relasi antar-tabel yang solid, meliputi:

1. **Data Penjualan:** Mencatat detail transaksi (ID, tanggal, kuantitas, harga, dan diskon).
2. **Data Produk:** Informasi katalog produk yang terintegrasi dengan kategori barang.
3. **Data Pelanggan:** Profil demografis entitas pelanggan berdasarkan lokasi geografis.
4. **Data Karyawan:** Informasi internal mengenai staf *sales* yang bertanggung jawab atas transaksi.

### 1.3 Tujuan Proyek

Tahap pengembangan ini bertujuan untuk:

1. Mengonstruksi arsitektur *Data Warehouse* berbasis **Star Schema**.
2. Mengeksekusi proses **ETL** (*Extract, Transform, Load*) menggunakan **Pentaho Data Integration** guna menjamin kebersihan dan integrasi data.
3. Melakukan migrasi data hasil transformasi ke database **MySQL** sebagai pondasi visualisasi.

## BAB 2

### PERANCANGAN DATA WAREHOUSE (STAR SCHEMA)

Sistem ini mengadopsi model **Star Schema** untuk menjamin efisiensi kueri. Struktur ini memusatkan metrik bisnis pada satu tabel fakta utama yang dikelilingi oleh tabel-tabel dimensi pendukung.

#### 2.1 Struktur Skema

1. **Fact Table (FactSales):** Berperan sebagai pusat penyimpanan transaksi dan metrik kuantitatif (*GrossAmount* dan *TotalAmount*).
2. **Dimension Tables:**
  - **DimCustomer:** Berisi atribut profil pelanggan.
  - **DimProduct:** Mencakup detail produk beserta kategorinya.
  - **DimEmployee:** Mendokumentasikan data karyawan pengelola pesanan.

## BAB 3

### IMPLEMENTASI ETL (EXTRACT, TRANSFORM, LOAD)

Dokumentasi alur kerja menggunakan **Pentaho Data Integration (PDI/Spoon)** adalah sebagai berikut:

#### 3.1 Implementasi Dimensi Produk (DimProduct)

Tahap ini menyatukan entitas produk dan kategori untuk memberikan konteks yang lebih jelas pada setiap item barang. Melalui fungsi *Stream Lookup*, data berhasil digabungkan sehingga produk seperti "Chai" secara otomatis terklasifikasi dalam kategori "Beverages".

##### Alur Transformasi:

1. **Input:** Mengambil data dari Input Categories dan Input Products.
2. **Stream Lookup:** Menggabungkan kedua aliran data berdasarkan CategoryID untuk mendapatkan CategoryName.
3. **Table Output:** Menyimpan hasil gabungan ke tabel dimproduct di database.

Gambar 3.1 Alur Transformasi Dimensi Produk di Pentaho

**Execution Results**

Logging Execution History Step Metrics Performance Graph Metrics Preview data

2025-12-11 12:39:38.563 - Spoon - Started the transformation execution.  
 2025-12-11 12:39:38.763 - dimproduct\_TA\_BI - Dispatching started for transformation [dimproduct\_TA\_BI]  
 2025-12-11 12:39:38.786 - Carte - Installing timer to purge stale objects after 1440 minutes.  
 2025-12-11 12:39:38.801 - Table output.0 - Connected to database [con\_db\_north] (commit=1000)  
 2025-12-11 12:39:38.850 - Input Categories.0 - Header row skipped in file 'D:\kuliah\smt5\BISNIS\_INTELIGENCE\archive (5)\categories.csv'  
 2025-12-11 12:39:38.851 - Input Products.0 - Header row skipped in file 'D:\kuliah\smt5\BISNIS\_INTELIGENCE\archive (5)\products.csv'  
 2025-12-11 12:39:38.852 - Input Categories.0 - Finished processing (I=9, O=0, R=0, W=8, U=0, E=0)  
 2025-12-11 12:39:38.853 - Input Products.0 - Finished processing (I=78, O=0, R=0, W=77, U=0, E=0)  
 2025-12-11 12:39:39.016 - Stream lookup.0 - Finished processing (I=0, O=0, R=85, W=77, U=0, E=0)  
 2025-12-11 12:39:39.467 - Table output.0 - Finished processing (I=0, O=77, R=77, W=77, U=0, E=0)  
 2025-12-11 12:39:39.468 - Spoon - The transformation has finished!!

**Hasil Eksekusi Data Produk:** Dapat dilihat pada *preview* data di bawah, data produk seperti "Chai" dan "Chang" berhasil digabungkan dengan kategori "Beverages".

Gambar 3.2 Preview Data Hasil Transformasi DimProduct

**Execution Results**

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	productID	productName	quantityPerUnit	unitPrice	discontinued	categoryID	CategoryName
1	1	Chai	10 boxes x 20 bags	18.0	0	1	Beverages
2	2	Chang	24 - 12 oz bottles	19.0	0	1	Beverages
3	3	Aniseed Syrup	12 - 550 ml bottles	10.0	0	2	Condiments
4	4	Chef Anton's Cajun Seasoning	48 - 6 oz jars	22.0	0	2	Condiments
5	5	Chef Anton's Gumbo Mix	36 boxes	21.35	1	2	Condiments
6	6	Grandma's Boysenberry Spread	12 - 8 oz jars	25.0	0	2	Condiments
7	7	Uncle Bob's Organic Dried Pears	12 - 1 lb pkgs.	30.0	0	7	Produce
8	8	Northwoods Cranberry Sauce	12 - 12 oz jars	40.0	0	2	Condiments
9	9	Mishi Kobe Niku	18 - 500 g pkgs.	97.0	1	6	Meat & Poultry

**Hasil Load di Database (MySQL):**

Gambar 3.3 Tabel dimproduct yang berhasil terisi di phpMyAdmin

Showing rows 0 - 24 (231 total, Query took 0.0007 seconds.)

`SELECT * FROM `dimproduct``

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

1 > >> ☐ Show all Number of rows: 25 Filter rows: Search this table Sort by key: None

Extra options

	ProductKey	ProductID	ProductName	CategoryName
<input type="checkbox"/> Edit Copy Delete	1	1	Chai	Beverages
<input type="checkbox"/> Edit Copy Delete	2	2	Chang	Beverages
<input type="checkbox"/> Edit Copy Delete	3	3	Aniseed Syrup	Condiments
<input type="checkbox"/> Edit Copy Delete	4	4	Chef Anton's Cajun Seasoning	Condiments
<input type="checkbox"/> Edit Copy Delete	5	5	Chef Anton's Gumbo Mix	Condiments
<input type="checkbox"/> Edit Copy Delete	6	6	Grandma's Boysenberry Spread	Condiments
<input type="checkbox"/> Edit Copy Delete	7	7	Uncle Bob's Organic Dried Pears	Produce
<input type="checkbox"/> Edit Copy Delete	8	8	Northwoods Cranberry Sauce	Condiments
<input type="checkbox"/> Edit Copy Delete	9	9	Mishi Kobe Niku	Meat & Poultry
<input type="checkbox"/> Edit Copy Delete	10	10	Ikura	Seafood
<input type="checkbox"/> Edit Copy Delete	11	11	Queso Cabrales	Dairy Products
<input type="checkbox"/> Edit Copy Delete	12	12	Queso Manchego La Pastora	Dairy Products

### 3.2 Implementasi Dimensi Pelanggan (DimCustomer)

Transformasi ini difokuskan pada pembersihan data master pelanggan untuk mendukung analisis distribusi pasar berdasarkan kota dan negara.

#### Alur Transformasi:

1. **Input Customers:** Membaca data mentah dari CSV.
2. **Table Output:** Memetakan dan memuat data bersih ke tabel dimcustomer di database.

Gambar 3.4 Alur Transformasi Dimensi Pelanggan

The screenshot shows the SAP Data Services Studio interface. At the top, there are tabs for 'Welcome!', 'dimproduct\_TA\_BI', 'DimCustomer\_TA\_BI...', 'DimEmployee\_TA\_BI...', and 'factsales\_TA\_BI'. The main workspace displays a transformation diagram with 'Input Customers' as the source and 'Table output' as the target, connected by a blue arrow. Below the workspace, the 'Execution Results' pane is open, showing a log of events. The log indicates that the transformation was successfully executed, with data being loaded into the 'Table output'.

**Execution Results**

- Logging
- Execution History
- Step Metrics
- Performance Graph
- Metrics
- Preview data

2025-12-11 12:40:30.029 - Spoon - Started the transformation execution.  
 2025-12-11 12:40:30.369 - Spoon - The transformation has finished!!  
 2025-12-11 12:41:24.302 - Spoon - Transformation opened.  
 2025-12-11 12:41:24.302 - Spoon - Launching transformation [DimCustomer\_TA\_BI]...  
 2025-12-11 12:41:24.302 - Spoon - Started the transformation execution.  
 2025-12-11 12:41:24.456 - DimCustomer\_TA\_BI - Dispatching started for transformation [DimCustomer\_TA\_BI]  
 2025-12-11 12:41:24.458 - Table output.0 - Connected to database [con\_db\_north] (commit=1000)  
 2025-12-11 12:41:24.482 - Input Customers.0 - Header row skipped in file 'file:///D:/kuliah/smt5/BISNIS\_INTELLIGENCE/archive%20(5)/customers.csv'  
 2025-12-11 12:41:24.483 - Input Customers.0 - Finished processing (I=92, O=0, R=0, W=91, U=0, E=0)  
 2025-12-11 12:41:24.573 - Table output.0 - Finished processing (I=0, O=91, R=91, W=91, U=0, E=0)  
 2025-12-11 12:41:24.573 - Spoon - The transformation has finished!!

**Hasil Load di Database (MySQL):** Tabel dimcustomer telah terisi dengan data pelanggan seperti "Alfreds Futterkiste" dari Berlin, Germany.

Gambar 3.5 Tabel dimcustomer di phpMyAdmin

The screenshot shows the phpMyAdmin interface for the 'dimcustomer' table. The table contains 12 rows of data, including 'Alfreds Futterkiste' from Berlin, Germany. The table structure is as follows:

CustomerKey	customerID	companyName	city	country
1	ALFKI	Alfreds Futterkiste	Berlin	Germany
2	ANATR	Ana Trujillo Emparedados y helados	Mexico City	Mexico
3	ANTON	Antonio Moreno Taqueria	Mexico City	Mexico
4	AROUT	Around the Horn	London	UK
5	BERGS	Berglunds snabbköp	Luleå	Sweden
6	BLAUS	Blauer See Delikatessen	Mannheim	Germany
7	BLONP	Blondesd's père et fils	Strasbourg	France
8	BOLID	Bólido Comidas preparadas	Madrid	Spain
9	BONAP	Bon app'	Marseille	France
10	BOTTM	Bottom-Dollar Markets	Tsawassen	Canada
11	BSBEV	B's Beverages	London	UK
12	CACTU	Cactus Comidas para llevar	Buenos Aires	Argentina

### 3.3 Implementasi Dimensi Karyawan (DimEmployee)

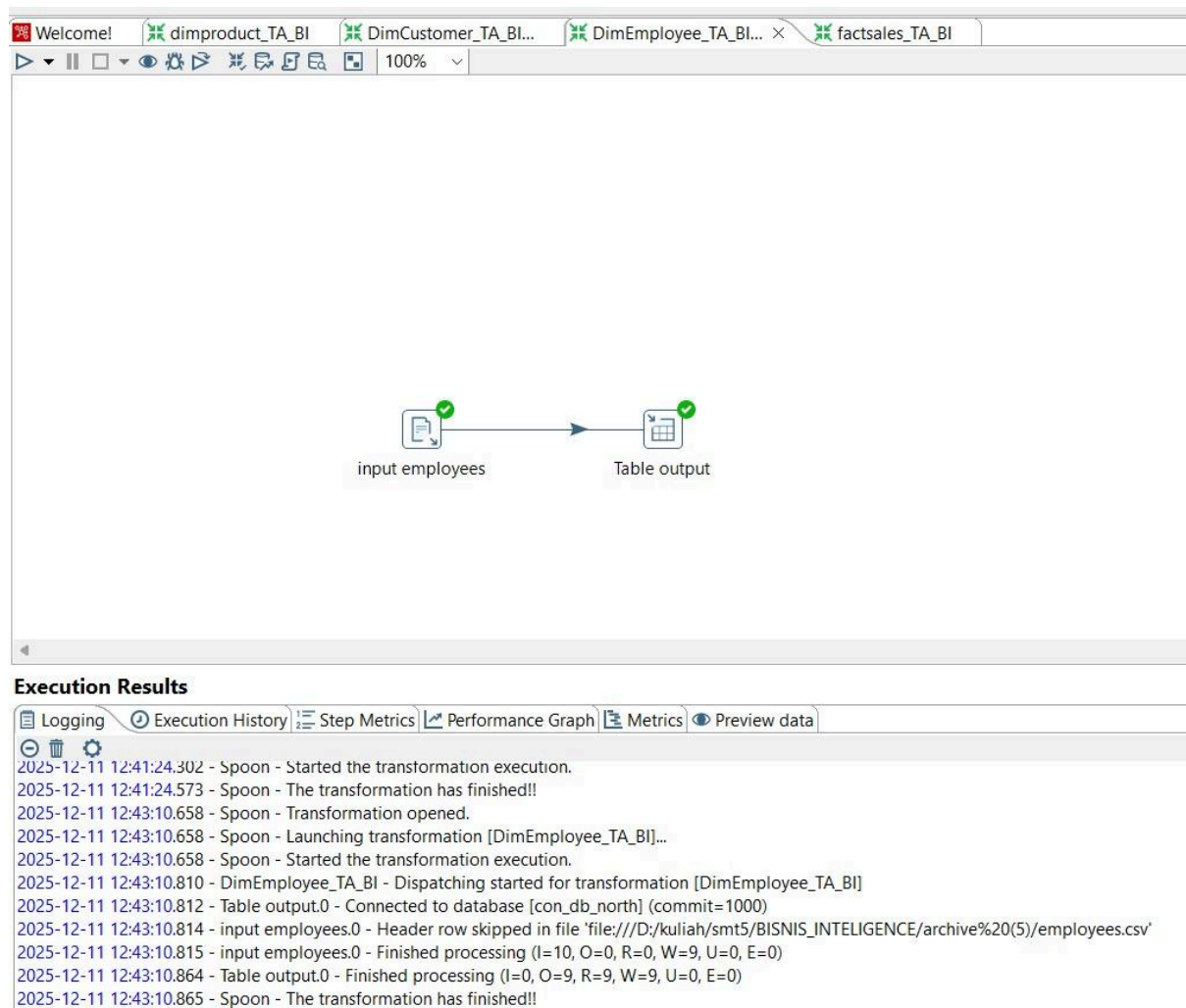


Proses ini menyiapkan data karyawan yang nantinya digunakan sebagai parameter evaluasi performa penjualan individu.

#### Alur Transformasi:

1. **Input Employees:** Mengambil data karyawan dari sumber.
2. **Table Output:** Memuat data ke tabel dimemployee.

Gambar 3.6 Alur Transformasi Dimensi Karyawan



Hasil Eksekusi Data Karyawan:

Gambar 3.7 Preview Data Karyawan (Nancy Davolio, Andrew Fuller, dll)

**Execution Results**

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	employeeID	employeeName	title	city	country	reportsTo
1	1	Nancy Davolio	Sales Representative	New York	USA	8
2	2	Andrew Fuller	Vice President Sales	New York	USA	<null>
3	3	Janet Leverling	Sales Representative	New York	USA	8
4	4	Margaret Peacock	Sales Representative	New York	USA	8
5	5	Steven Buchanan	Sales Manager	London	UK	2
6	6	Michael Suyama	Sales Representative	London	UK	5
7	7	Robert King	Sales Representative	London	UK	5
8	8	Laura Callahan	Sales Manager	New York	USA	2
9	9	Anne Dodsworth	Sales Representative	London	UK	5

Hasil Load di Database (MySQL):

Gambar 3.8 Tabel dimemployee di phpMyAdmin

EmployeeKey	EmployeeID	EmployeeName	Title
1	1	Nancy Davolio	Sales Representative
2	2	Andrew Fuller	Vice President Sales
3	3	Janet Leverling	Sales Representative
4	4	Margaret Peacock	Sales Representative
5	5	Steven Buchanan	Sales Manager
6	6	Michael Suyama	Sales Representative
7	7	Robert King	Sales Representative
8	8	Laura Callahan	Sales Manager
9	9	Anne Dodsworth	Sales Representative

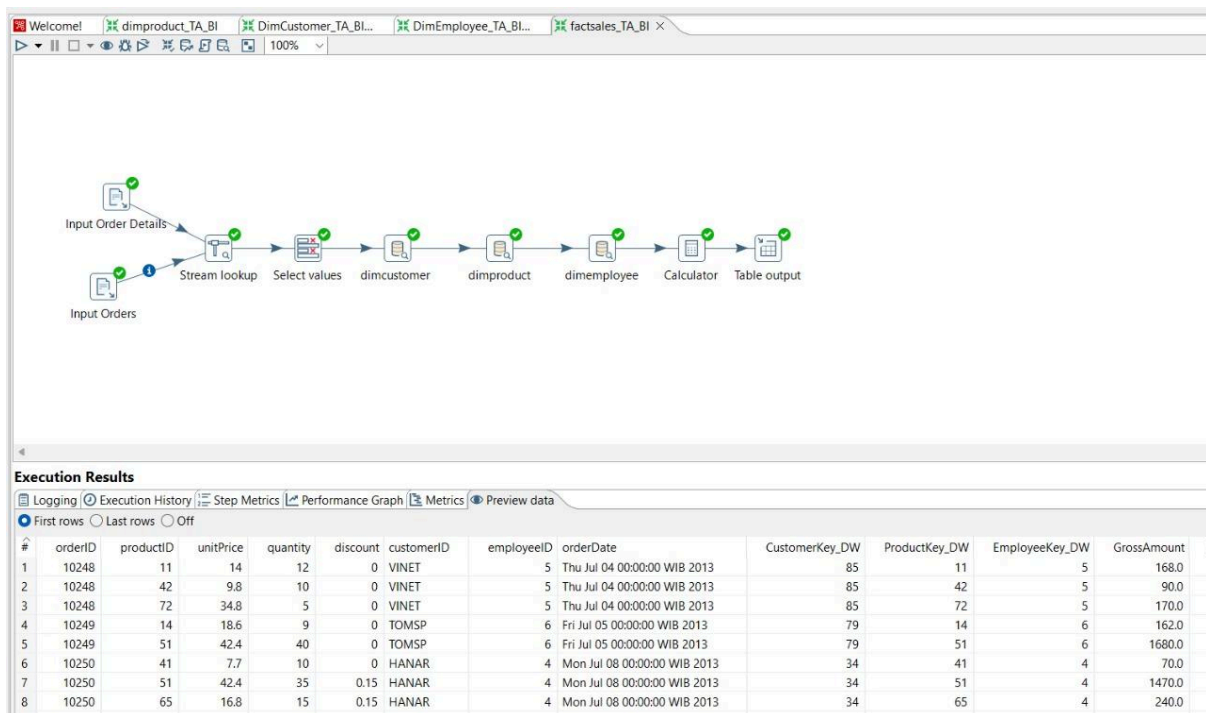
### 3.4 Implementasi Tabel Fakta (FactSales)

Sebagai tahap paling kompleks, proses ini melibatkan penggabungan header dan detail pesanan, pencarian kunci dimensi (*lookup*), serta kalkulasi finansial. Metrik yang dihitung meliputi:

#### Alur Transformasi:

1. **Input Data:** Mengambil data dari orders.csv dan order\_details.csv.
2. **Stream Lookup:** Menggabungkan detail pesanan dengan header pesanan berdasarkan OrderID.
3. **Dimension Lookups (Database Lookup):**
  - Mencari CustomerKey dari tabel dimcustomer.
  - Mencari ProductKey dari tabel dimproduct.
  - Mencari EmployeeKey dari tabel dimemployee.
4. **Calculator:** Melakukan perhitungan matematis untuk kolom *Measure*:
  - $GrossAmount = UnitPrice * Quantity$
  - $TotalAmount = GrossAmount - (GrossAmount * Discount)$
5. **Table Output:** Menyimpan hasil akhir ke tabel factsales.

Gambar 3.9 Alur Transformasi Lengkap FactSales



**Hasil Perhitungan (Calculator & Lookups):** Pada gambar di bawah, terlihat kolom CustomerKey\_DW, ProductKey\_DW, serta hasil perhitungan GrossAmount dan TotalAmount sudah terbentuk.

Gambar 3.10 Preview Data FactSales dengan Key dan Measure yang sudah dihitung

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

unitPrice	quantity	discount	customerID	employeeID	orderDate	CustomerKey_DW	ProductKey_DW	EmployeeKey_DW	GrossAmount	DiscountAmount	TotalAmount
14	12	0	VINET	5	Thu Jul 04 00:00:00 WIB 2013	85	11	5	168.0	0.0	168.0
9.8	10	0	VINET	5	Thu Jul 04 00:00:00 WIB 2013	85	42	5	90.0	0.0	90.0
34.8	5	0	VINET	5	Thu Jul 04 00:00:00 WIB 2013	85	72	5	170.0	0.0	170.0
18.6	9	0	TOMSP	6	Fri Jul 05 00:00:00 WIB 2013	79	14	6	162.0	0.0	162.0
42.4	40	0	TOMSP	6	Fri Jul 05 00:00:00 WIB 2013	79	51	6	1680.0	0.0	1680.0
7.7	10	0	HANAR	4	Mon Jul 08 00:00:00 WIB 2013	34	41	4	70.0	0.0	70.0
42.4	35	0.15	HANAR	4	Mon Jul 08 00:00:00 WIB 2013	34	51	4	1470.0	220.5	1249.5
16.8	15	0.15	HANAR	4	Mon Jul 08 00:00:00 WIB 2013	34	65	4	240.0	36.0	204.0

**Hasil Load di Database (MySQL):** Tabel fakta berhasil terisi dan siap digunakan untuk analisis/visualisasi.

Gambar 3.11 Tabel factsales di phpMyAdmin



## BAB 4

### ANALISIS DAN VISUALISASI DASHBOARD

Berdasarkan data yang telah diproses ke dalam database MySQL, berikut adalah hasil visualisasi KPI Sales:

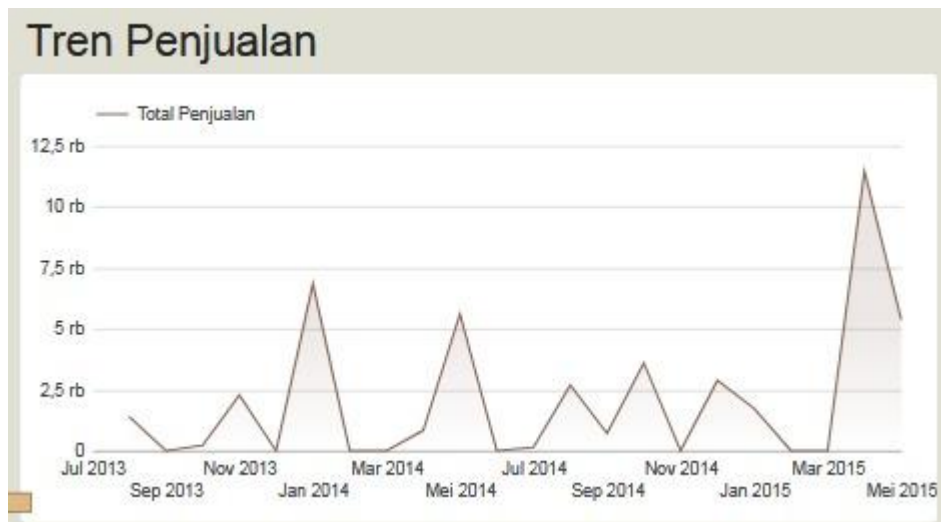
#### 4.1 Ringkasan Pendapatan Kotor



Total pendapatan kotor yang diperoleh perusahaan selama periode pengamatan (Juli 2013 - Mei 2015) mencapai 1.252.483. Nilai ini menunjukkan akumulasi keseluruhan transaksi penjualan dari berbagai kategori produk dan wilayah pemasaran.

Selama periode Juli 2013 hingga Mei 2015, Northwind Traders berhasil membukukan total pendapatan kotor sebesar **1.252.482,77**. Angka ini merupakan representasi dari akumulasi transaksi global di seluruh kategori produk

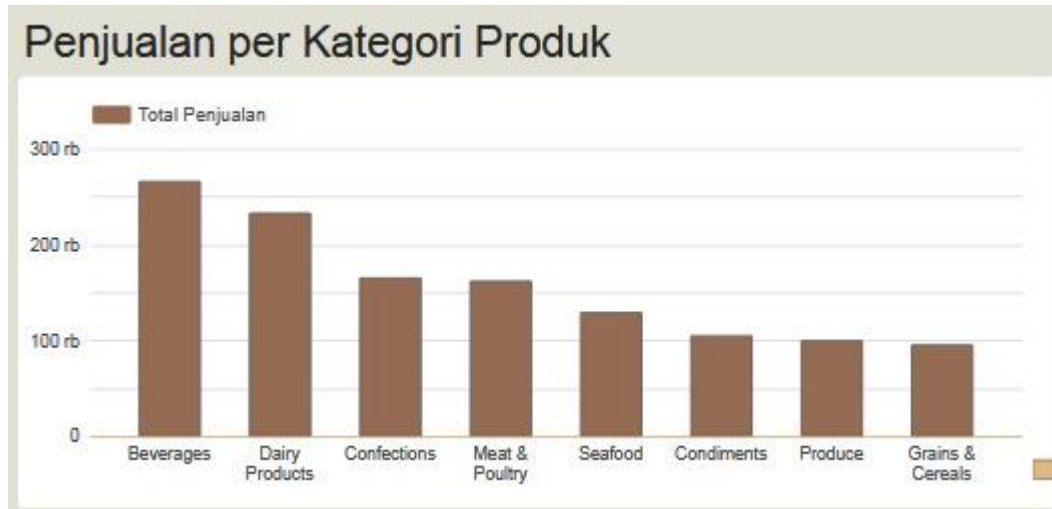
#### 4.2 Analisis Tren Penjualan



Visualisasi tren penjualan menunjukkan fluktuasi nilai dari waktu ke waktu, dengan kecenderungan meningkat pada periode-periode tertentu dan spesifik mengalami peningkatan tajam di sekitar Maret - April 2015. Pola ini mengindikasikan adanya pengaruh musiman serta dinamika permintaan pasar terhadap volume penjualan.

Grafik tren menunjukkan fluktuasi yang dinamis dengan indikasi pertumbuhan positif. Lonjakan volume penjualan yang paling signifikan terjadi pada periode **Maret hingga April 2015**, yang mencerminkan adanya pengaruh musiman serta perubahan permintaan pasar secara tajam.

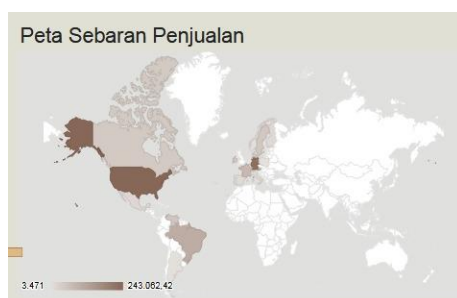
#### 4.3 Kontribusi per Kategori Produk



Penjualan berdasarkan kategori produk memperlihatkan kontribusi yang berbeda-beda terhadap total pendapatan perusahaan. Beberapa kategori seperti **Beverages**, **Dairy Products**, dan **Confections** memberikan kontribusi penjualan yang lebih dominan dibandingkan kategori lainnya. Hal ini menunjukkan bahwa produk-produk tersebut memiliki tingkat permintaan yang tinggi dan menjadi penggerak utama pendapatan perusahaan.

Analisis kategori menunjukkan bahwa **Beverages (Minuman)**, **Dairy Products (Produk Susu)**, dan **Confections (Manisan)** adalah penggerak utama pendapatan perusahaan. Ketiga kategori ini memiliki tingkat permintaan yang jauh lebih dominan dibandingkan kategori lainnya.

#### 4.4 Sebaran Geografis Penjualan





Peta sebaran penjualan menunjukkan distribusi transaksi yang tersebar di berbagai negara dengan tingkat penjualan yang bervariasi. Wilayah dengan intensitas warna lebih tinggi merepresentasikan kontribusi penjualan yang lebih besar, yang mana didominasi oleh **USA** dan **Germany**. Visualisasi ini membantu mengidentifikasi pasar potensial serta wilayah strategis yang berperan penting dalam mendukung kinerja penjualan global Northwind Traders sebagai penyuplai makanan mewah (*gourmet*).

Penjualan tersebar secara global dengan dominasi pasar yang kuat di wilayah **Amerika Serikat (USA)** dan **Jerman**. Hal ini mengukuhkan posisi Northwind Traders sebagai pemasok utama produk makanan mewah (*gourmet*) di wilayah strategis tersebut.

#### 4.5 Detail Operasional Sales

	Tanggal	Kategori	Customer	Negara	Produk	Qty	Pegawai	Total Penjualan
1.	4 Jul 2013	Dairy Products	Vins et alcools Chevalier	France	Queso Cabrales	12	Steven Buchanan	168
2.	4 Jul 2013	Grains & Cereals	Vins et alcools Chevalier	France	Singaporean Hokkien Fried Mee	10	Steven Buchanan	90
3.	4 Jul 2013	Dairy Products	Vins et alcools Chevalier	France	Mozzarella di Giovanni	5	Steven Buchanan	170
4.	5 Jul 2013	Produce	Toms Spezialitäten	Germany	Tofu	9	Michael Suyama	162
5.	5 Jul 2013	Produce	Toms Spezialitäten	Germany	Manjimup Dried Apples	40	Michael Suyama	1.680
6.	8 Jul 2013	Seafood	Hanari Carnes	Brazil	Jack's New England Clam Cho...	10	Margaret Peacock	70
7.	8 Jul 2013	Produce	Hanari Carnes	Brazil	Manjimup Dried Apples	35	Margaret Peacock	1.249,5
8.	8 Jul 2013	Condiments	Hanari Carnes	Brazil	Louisiana Fiery Hot Pepper Sauce	15	Margaret Peacock	204
9.	8 Jul 2013	Grains & Cereals	Victuailles en stock	France	Gustaf's Knackbröd	6	Janet Leverling	91,2
10.	8 Jul 2013	Grains & Cereals	Victuailles en stock	France	Ravioli Angelo	15	Janet Leverling	213,75
11.	8 Jul 2013	Condiments	Victuailles en stock	France	Louisiana Fiery Hot Pepper Sauce	20	Janet Leverling	220

Dashboard juga menyediakan data granular yang mencakup keterlibatan staf *sales* (seperti Margaret Peacock atau Steven Buchanan) dalam setiap transaksi, memberikan transparansi penuh terhadap operasional penjualan harian.

## **BAB 5**

### **KESIMPULAN**

Melalui implementasi ETL dan perancangan *Data Warehouse* yang sistematis, data mentah Northwind Traders telah berhasil ditransformasi menjadi aset informasi strategis. Hasil analisis menyimpulkan bahwa perusahaan memiliki fundamental yang kuat di pasar Barat (USA & Jerman), dengan kategori *Beverages* sebagai kontributor utama dalam struktur pendapatan perusahaan.