

**LAPORAN FINAL PROJECT**  
**PENGEMBANGAN DATA WAREHOUSE DAN OLAP STUDI KASUS**  
**DVD RENTAL**  
**SEMESTER GANJIL**  
**2024/2025**

**Dosen Pengampu**  
Mohamad Irwan Afandi, ST., MSC.



**Disusun Oleh**  
Hilmi Arya Rafwa Muhammad (22082010061)

**PROGRAM STUDI SISTEM INFORMASI**  
**FAKULTAS ILMU KOMPUTER**  
**Universitas Pembangunan Nasional "Veteran" Jawa Timur**  
**S U R A B A Y A**  
**2024**

## DAFTAR ISI

<b>DAFTAR ISI.....</b>	<b>1</b>
<b>DAFTAR GAMBAR.....</b>	<b>2</b>
<b>BAB I PENDAHULUAN.....</b>	<b>3</b>
1.1 Latar Belakang.....	3
1.2 Rumusan Masalah.....	3
1.3 Tujuan Kegiatan.....	3
1.4 Batasan.....	4
<b>BAB II METODOLOGI.....</b>	<b>5</b>
2.1 Daftar Pustaka.....	5
2.1.1 DVD Rental (Studi Kasus).....	5
2.1.2 PostgreSQL.....	5
2.1.3 DBeaver.....	5
2.1.4 Pentaho Data Integration (PDI).....	6
2.1.5 XAMPP.....	6
2.1.6 MySQL.....	6
2.1.7 Visual Studio Code.....	6
2.1.8 Mondrian.....	6
2.1.9 Tomcat.....	7
2.2 Metode Pengembangan.....	7
2.2.1 Perancangan Skema Data Warehouse.....	7
2.2.2 Ekstraksi Data ke MySql.....	7
2.2.3 Implementasi CUBE dan OLAP.....	7
<b>BAB III IMPLEMENTASI DAN HASIL.....</b>	<b>9</b>
3.1 Perancangan Skema Data Warehouse.....	9
3.3.1 Membuat Database dan Schema.....	9
3.3.2 Membuat Tabel Dimensi.....	9
3.3.3 Membuat Tabel Fakta.....	15
3.3.4 Validasi dan Penyimpanan.....	16
3.2 Ekstraksi Data ke MySQL.....	16
3.3 Implementasi CUBE dan OLAP.....	17
3.3.1 Query.....	17
3.3.2 Hasil.....	19
<b>BAB IV PENUTUPAN.....</b>	<b>23</b>
4.1 Kesimpulan.....	23
4.2 Saran.....	23
<b>LAMPIRAN.....</b>	<b>24</b>

## DAFTAR GAMBAR

Gambar 3.1 Tampilan Transformasi tf_dim_customer.....	11
Gambar 3.2 Data Customer pada Tabel dim_customer.....	11
Gambar 3.3 Tampilan Transformasi tf_dim_film.....	12
Gambar 3.4 Data Film pada Tabel dim_film.....	12
Gambar 3.5 Tampilan Transformasi tf_dim_staff.....	13
Gambar 3.6 Data Staff pada Tabel dim_staff.....	13
Gambar 3.7 Tampilan Transformasi tf_dim_store.....	13
Gambar 3.8 Data Store pada Tabel dim_store.....	14
Gambar 3.9 Tampilan Transformasi tf_dim_paymentdate.....	14
Gambar 3.10 Tampilan Transformasi tf_dim_rentaldate.....	15
Gambar 3.11 Tampilan Transformasi tf_dim_returndate.....	15
Gambar 3.12 Data Film pada Tabel dim_paymentdate.....	15
Gambar 3.13 Data Film pada Tabel dim_rentaldate.....	16
Gambar 3.14 Data Film pada Tabel dim_returndate.....	16
Gambar 3.15 Tampilan Transformasi tf_fact_sales.....	17
Gambar 3.16 Data Sales pada Tabel fact_sales.....	17
Gambar 3.19 Tampilan Awal Mondrian.....	20
Gambar 3.20 Tampilan Dimensi Store Mondrian.....	21
Gambar 3.21 Tampilan Dimensi Customer Country Mondrian.....	21
Gambar 3.22 Tampilan Dimensi Customer District Mondrian.....	22
Gambar 3.23 Tampilan Dimensi Customer City Mondrian.....	22
Gambar 3.24 Tampilan Dimensi Film Category Mondrian.....	23

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Dalam era digital yang semakin maju, pengelolaan data menjadi salah satu aspek penting dalam pengambilan keputusan bisnis. Industri penyewaan DVD, meskipun terlihat tradisional, masih menyimpan banyak peluang dalam memaksimalkan efisiensi operasional dan meningkatkan profitabilitas melalui analisis data yang tepat. Database DVD Rental menyediakan kumpulan data yang komprehensif terkait pelanggan, transaksi, kategori film, dan performa toko. Dengan memanfaatkan konsep Data Warehouse dan OLAP (Online Analytical Processing), analisis mendalam dapat dilakukan untuk menemukan pola perilaku pelanggan, tren penyewaan, serta performa toko berdasarkan lokasi geografis dan kategori film.

Proyek ini bertujuan untuk merancang dan mengimplementasikan Data Warehouse menggunakan alat seperti PostgreSQL, DBeaver, dan Pentaho Data Integration (PDI) serta memanfaatkan Mondrian OLAP untuk analisis multidimensi guna memberikan insight yang lebih baik bagi pengambilan keputusan bisnis.

### **1.2 Rumusan Masalah**

- 1.) Bagaimana merancang skema Data Warehouse untuk studi kasus DVD Rental?
- 2.) Bagaimana melakukan proses ETL (Extract, Transform, Load) yang efektif menggunakan Pentaho Data Integration?
- 3.) Bagaimana mengimplementasikan analisis data menggunakan OLAP (Mondrian OLAP) untuk mengevaluasi performa penyewaan DVD berdasarkan kategori film, lokasi toko, dan demografi pelanggan?
- 4.) Bagaimana memvisualisasikan hasil analisis data untuk mendukung pengambilan keputusan strategis?

### **1.3 Tujuan Kegiatan**

- 1.) Merancang skema Data Warehouse dengan pendekatan Star Schema yang mencakup tabel dimensi dan fakta.
- 2.) Melakukan proses ETL (Extract, Transform, Load) untuk mempersiapkan data dari database operasional ke dalam Data Warehouse.
- 3.) Mengimplementasikan OLAP Cube untuk memungkinkan analisis data multidimensi.

- 4.) Menyediakan insight yang dapat diandalkan untuk pengambilan keputusan terkait operasional penyewaan DVD.

#### **1.4 Batasan**

- 1.) Studi kasus terbatas pada dataset DVD Rental.
- 2.) Alat yang digunakan terbatas pada PostgreSQL, DBeaver, Pentaho Data Integration (PDI), Visual Studio Code, dan Mondrian OLAP.
- 3.) Analisis data hanya difokuskan pada tiga aspek utama: Performa Toko, Wilayah Tempat Tinggal Pelanggan, dan Kategori Film.
- 4.) Visualisasi data hanya dilakukan menggunakan antarmuka yang disediakan oleh Mondrian OLAP.

## **BAB II**

### **METODOLOGI**

#### **2.1 Daftar Pustaka**

##### **2.1.1 DVD Rental (Studi Kasus)**

Studi kasus yang diangkat dalam kegiatan praktik ini berfokus pada Analisis Hasil Penyewaan DVD Berdasarkan Performa Toko, Wilayah Tempat Tinggal Customer, dan Kategori Film pada Database DVD Rental. Database ini merupakan salah satu contoh dataset yang umum digunakan untuk mempelajari sistem manajemen basis data relasional (RDBMS) dan analisis data dalam skenario bisnis penyewaan DVD. Database DVD Rental menyediakan berbagai tabel yang mencakup informasi pelanggan, inventaris film, transaksi penyewaan, pembayaran, serta kategori film.

Secara umum, database DVD Rental dirancang untuk merepresentasikan operasi bisnis dalam industri penyewaan DVD. Tabel-tabel utama yang digunakan dalam studi kasus ini meliputi Tabel Customer, Film, Rental, Inventory, Store, Payment, dan Tabel Category.

Dalam konteks studi kasus ini, fokus analisis diarahkan pada bagaimana pola penyewaan DVD dapat diidentifikasi dan dimanfaatkan untuk meningkatkan efisiensi operasional serta merancang strategi pemasaran yang lebih efektif. Hal ini melibatkan pemahaman mendalam tentang Perilaku Pelanggan, Performa Toko, Wilayah Tempat Tinggal Customer, dan Kategori Film.

##### **2.1.2 PostgreSQL**

PostgreSQL adalah object-relational database management system (ORDBMS) yang dikembangkan di Departemen Ilmu Komputer Universitas California di Berkeley. PostgreSQL mendukung sebagian besar standar SQL dan menawarkan banyak fitur modern, seperti complex queries, foreign keys, dan triggers. PostgreSQL digunakan karena fleksibel untuk menangani data dengan struktur yang kompleks dan kemampuan integrasi dengan berbagai tools ETL dan visualisasi.

##### **2.1.3 DBeaver**

DBeaver adalah alat manajemen database universal yang mendukung berbagai sistem RDBMS, termasuk PostgreSQL. Alat ini menyediakan UI yang ramah pengguna untuk menjalankan query, memvisualisasikan skema melalui ERD, dan mengelola data. DBeaver dipilih karena mudah dalam eksplorasi database, fitur visualisasi skema yang

mendukung analisis data, dan mendukung PostgreSQL yang mempermudah pengelolaan data.

#### **2.1.4 Pentaho Data Integration (PDI)**

Pentaho Data Integration adalah alat untuk proses ETL (Extract, Transform, Load) yang memadukan berbagai set data menjadi satu sumber sebagai dasar untuk kebutuhan analisis. Dikelola dengan grafis interface drag-and-drop sehingga memudahkan untuk melacak asal data, tujuan data, dan cara transformasinya. Pentaho dipilih untuk proses ETL karena mampu menangani berbagai format data, fitur transformasi yang bagus, dan penggunaan yang mudah dengan interface drag-and-drop.

#### **2.1.5 XAMPP**

XAMPP adalah perangkat lunak open-source yang menyediakan paket solusi lengkap untuk pengembangan aplikasi berbasis web. Paket ini mencakup Apache sebagai web server, MySQL sebagai sistem manajemen basis data, dan berbagai tool pendukung lainnya seperti PHP dan Perl. XAMPP dipilih dalam proyek ini untuk memfasilitasi migrasi skema Data Warehouse dari PostgreSQL ke MySQL serta untuk menjalankan Mondrian OLAP menggunakan server Tomcat yang disertakan di dalamnya.

#### **2.1.6 MySQL**

MySQL adalah salah satu sistem manajemen basis data relasional (RDBMS) yang paling populer di dunia. Dikembangkan oleh Oracle, MySQL terkenal karena kecepatan, keandalan, dan kemudahan penggunaannya. Dalam proyek ini, MySQL digunakan untuk menyimpan skema Data Warehouse setelah proses migrasi dari PostgreSQL, serta sebagai sumber data untuk analisis OLAP.

#### **2.1.7 Visual Studio Code**

Visual Studio Code (VS Code) adalah editor teks ringan yang dikembangkan oleh Microsoft. VS Code mendukung berbagai bahasa pemrograman dan dilengkapi dengan ekosistem ekstensi yang luas. Dalam proyek ini, Visual Studio Code digunakan untuk mengedit file konfigurasi OLAP seperti fact\_sales.xml dan fact\_sales.jsp dengan bantuan ekstensi pendukung.

#### **2.1.8 Mondrian**

Mondrian adalah server OLAP open-source yang memungkinkan analisis data multidimensi secara efisien. Dalam proyek ini, Mondrian digunakan untuk mengimplementasikan cube OLAP yang terhubung ke Data Warehouse MySQL.

Dengan Mondrian, pengguna dapat mengeksplorasi data melalui berbagai hierarki dimensi untuk mendapatkan insight yang bermanfaat.

#### **2.1.9 Tomcat**

Apache Tomcat adalah server aplikasi open-source untuk menjalankan aplikasi web berbasis Java. Tomcat digunakan dalam proyek ini untuk menjalankan Mondrian OLAP dan menyajikan antarmuka analisis data yang diakses melalui browser web.

### **2.2 Metode Pengembangan**

#### **2.2.1 Perancangan Skema Data Warehouse**

Tahap Perancangan Skema Data Warehouse berfokus pada penyusunan dan pengorganisasian data ke dalam struktur yang lebih terorganisir dan siap untuk dianalisis. Proses dimulai dengan pembuatan database `dwh_dvdrental` di DBeaver dan dilanjutkan dengan pembuatan skema `dwh` di dalam database tersebut. Skema ini berfungsi sebagai wadah utama untuk tabel dimensi dan tabel fakta. Tabel dimensi yang dibuat meliputi tabel `customers`, `film`, `staff`, `store`, `payment date`, `rental date`, dan `return date`. Setiap tabel dimensi dirancang untuk menyimpan atribut yang mendeskripsikan entitas terkait, sementara tabel fakta, seperti `sales`, berisi data transaksi atau aktivitas utama yang akan dianalisis. Proses transformasi data dilakukan dengan cermat untuk memastikan konsistensi dan keakuratan data. Hasil dari tahap ini adalah struktur Data Warehouse yang siap digunakan untuk penyusunan Datamart.

#### **2.2.2 Ekstraksi Data ke MySQL**

Proses ekstraksi data ke MySQL dilakukan untuk memastikan bahwa skema Data Warehouse yang dikembangkan dapat digunakan dengan optimal dalam implementasi OLAP. Langkah pertama adalah melakukan backup pada database `dwh_dvdrental` menggunakan DBeaver, di mana seluruh tabel dan skema diekspor ke dalam file SQL. Setelah itu, layanan MySQL diaktifkan melalui XAMPP, dan sebuah database baru bernama "`dwh`" dibuat menggunakan phpMyAdmin. File SQL hasil backup kemudian diunggah ke phpMyAdmin dan query-nya dijalankan untuk mengimpor data ke database MySQL. Langkah terakhir adalah memvalidasi struktur data menggunakan phpMyAdmin untuk memastikan bahwa semua tabel, skema, dan data telah terimpor dengan benar tanpa adanya kesalahan.

#### **2.2.3 Implementasi CUBE dan OLAP**

Implementasi cube dan OLAP dimulai dengan pembuatan file konfigurasi bernama `fact_sales.xml` yang mendefinisikan struktur cube OLAP. File ini mencakup



berbagai dimensi seperti Dimensi Store, Dimensi Customer Demographics, dan Dimensi Film Category, serta measure seperti Total Sales dan Total Amount untuk analisis data. Selanjutnya, dibuat file `fact_sales.jsp` untuk menghubungkan skema cube dengan database MySQL menggunakan JDBC. File ini juga berisi query MDX untuk mengambil data dari cube OLAP dan menampilkan hasil analisis dalam antarmuka web. Setelah file konfigurasi selesai, kedua file ini ditempatkan di direktori `webapps` pada Tomcat. Server Tomcat kemudian dijalankan melalui XAMPP, dan antarmuka OLAP dapat diakses melalui browser untuk memverifikasi bahwa cube berfungsi dengan baik. Validasi dan pengujian dilakukan untuk memastikan bahwa semua dimensi, measure, dan query MDX bekerja sesuai dengan yang diharapkan.

## **BAB III**

### **IMPLEMENTASI DAN HASIL**

#### **3.1 Perancangan Skema Data Warehouse**

Desain data warehouse merupakan langkah krusial dalam proses pengolahan data. Pada tahap ini, data dari staging area dimodelkan menggunakan pendekatan star schema, yang terdiri dari tabel dimensi dan tabel fakta. Pendekatan ini dipilih karena kesederhanaannya, kemudahan dalam interpretasi, dan efisiensinya untuk query analitik. Berikut adalah langkah-langkah desain data warehouse untuk dvdrental:

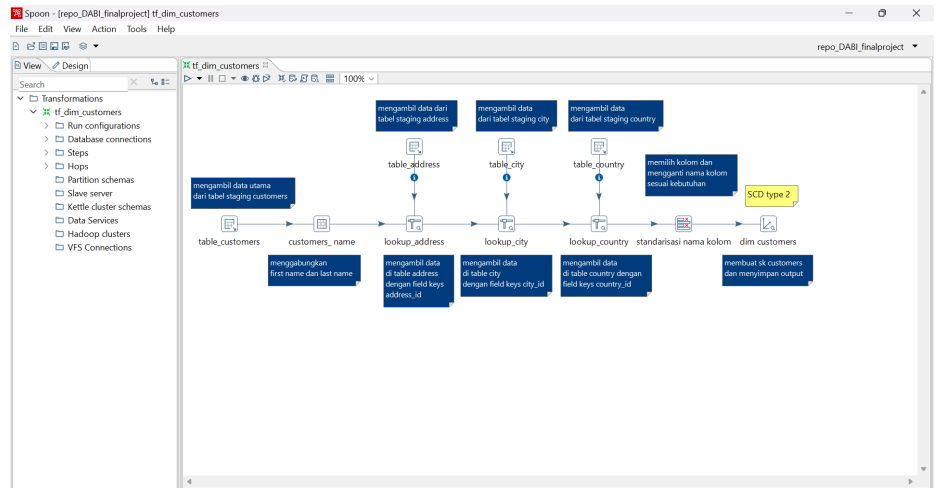
##### **3.3.1 Membuat Database dan Schema**

- Buat database baru dengan nama `dwh_dvdrental` menggunakan DBeaver.
- Di dalam database tersebut, buat schema baru dengan nama `dwh`, yang akan digunakan untuk menyimpan tabel-tabel dimensi dan fakta.

##### **3.3.2 Membuat Tabel Dimensi**

Tabel dimensi berisi atribut deskriptif yang membantu memberikan konteks pada data. Untuk database `dvdrental`, tabel dimensi yang dirancang meliputi:

- Dimensi Customers
  - Bertujuan untuk memahami dan mengelompokkan data pelanggan berdasarkan atribut tertentu, seperti nama, alamat, dan status aktif.
  - Data diambil dari tabel staging seperti `customer`, `address`, `city`, dan `country`.
  - Transformasi melibatkan penggabungan nama depan dan belakang menjadi `customer_name`, serta penyesuaian nama kolom untuk konsistensi (misalnya `email` menjadi `customer_email`).
  - Output disimpan di tabel `dim_customer` dengan penerapan SCD Type 2 untuk menangani perubahan historis data pelanggan.
  - Gambar 3.1 dan 3.2 berikut merupakan tampilan dari transformasi `tf_dim_customer` dan hasil data ditampung di tabel `dim_customer`.



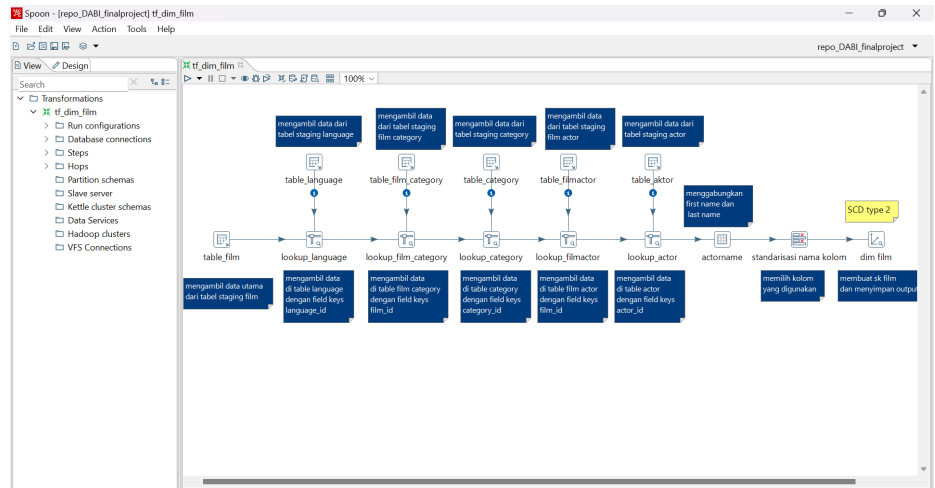
**Gambar 3.1** Tampilan Transformasi tf\_dim\_customer

id	customer_id	date	customer_name	customer_email	customer_phone	customer_address	customer_district
1	1	1900-01-01 00:00:00.000	Jared Ty	jared.ty@baklaccustomer.org	3553111597	100 Oakhurst Street	West Ave
2	2	1900-01-01 00:00:00.000	Mary Smith	mary.smith@baklaccustomer.org	2332134569	103 Farrow Way	San Francisco
3	3	1900-01-01 00:00:00.000	Patricia Johnson	patricia.johnson@baklaccustomer.org	83861528648	1121 Loop Avenue	California
4	4	1900-01-01 00:00:00.000	Linda Williams	linda.williams@baklaccustomer.org	44887790408	692 Joliet Street	Alaska
5	5	1900-01-01 00:00:00.000	Bethany Jones	bethany.jones@baklaccustomer.org	70524002527	1560 Ingot Mines	Mississippi
6	6	1900-01-01 00:00:00.000	Elizabeth Brown	elizabeth.brown@baklaccustomer.org	1506160674	51 Idyll Parkway	Nantux
7	7	1900-01-01 00:00:00.000	Jennifer Davis	jennifer.davis@baklaccustomer.org	88826262444	1795 Santiago de Compostela Way	Texas
8	8	1900-01-01 00:00:00.000	Maria Miller	maria.miller@baklaccustomer.org	71827122073	900 Seneca de Compostela Parkway	Central Serbia
9	9	1900-01-01 00:00:00.000	Susan Wilson	susan.wilson@baklaccustomer.org	67282323979	478 Joliet Drive	Hamilton
10	10	1900-01-01 00:00:00.000	Margaret Moore	margaret.moore@baklaccustomer.org	38865762348	611 Kinsey Way	Manjot
11	11	1900-01-01 00:00:00.000	Dorothy Taylor	dorothy.taylor@baklaccustomer.org	14880881819	1151 Sal Drive	Edgah
12	12	1900-01-01 00:00:00.000	Lisa Anderson	lisa.anderson@baklaccustomer.org	61529727745	1542 Tardac Parkway	Kanagawa
13	13	1900-01-01 00:00:00.000	Nancy Thomas	nancy.thomas@baklaccustomer.org	46188188014	270 Ansonia Parkway	Mayaguez
14	14	1900-01-01 00:00:00.000	Renee Jackson	renee.jackson@baklaccustomer.org	89547988718	270 Ansonia Parkway	Omaha
15	15	1900-01-01 00:00:00.000	Betty White	betty.white@baklaccustomer.org	51738181423	778 Bygonia Avenue	California
16	16	1900-01-01 00:00:00.000	Robin King	robin.king@baklaccustomer.org	98881182254	410 Regan Lane	Madhya Pradesh
17	17	1900-01-01 00:00:00.000	Sandra Martin	sandra.martin@baklaccustomer.org	98881123197	369 Toulouse Parkway	England
18	18	1900-01-01 00:00:00.000	Donna Thompson	donna.thompson@baklaccustomer.org	40773241482	270 Toulon Boulevard	Kashyria
19	19	1900-01-01 00:00:00.000	Carol Garcia	carol.garcia@baklaccustomer.org	14781954888	320 Bird Avenue	Kaduna
20	20	1900-01-01 00:00:00.000	Ruth Martinez	ruth.martinez@baklaccustomer.org	27272357893	1417 Lander Avenue	Northern Cape
21	21	1900-01-01 00:00:00.000	Sharon Robinson	sharon.robinson@baklaccustomer.org	14483388132	1688 Okara Way	Madhya Pradesh
22	22	1900-01-01 00:00:00.000	Michelle Clark	michelle.clark@baklaccustomer.org	88273730083	282 A Grove La Grand Parkway	Dhar
23	23	1900-01-01 00:00:00.000	Laura Rodriguez	laura.rodriguez@baklaccustomer.org	16186837433	28 Charlotte Avenue Street	Rabat Sal Zammour-2
24	24	1900-01-01 00:00:00.000	Sarah Lewis	sarah.lewis@baklaccustomer.org	80273123833	1780 Hono Boulevard	Leppia
25	25	1900-01-01 00:00:00.000	Kimberly Lee	kimberly.lee@baklaccustomer.org	15420181242	96 Telara Way	Cebu
26	26	1900-01-01 00:00:00.000	Deborah Walker	deborah.walker@baklaccustomer.org	19489494706	934 San Felipe de Puerto Plata Street	Sind
27	27	1900-01-01 00:00:00.000	Jessica Hall	jessica.hall@baklaccustomer.org	89888977782	18 Dunbar Boulevard	Cile de
28	28	1900-01-01 00:00:00.000	Shirley Allen	shirley.allen@baklaccustomer.org	66256770122	121 Bethesda Place	Southern Mindanao
29	29	1900-01-01 00:00:00.000	Madara King	madara.king@baklaccustomer.org	52525548078	1405 Ansonia Manor	Bihar
30	30	1900-01-01 00:00:00.000	Cynthia Young	cynthia.young@baklaccustomer.org	67820867005	33 Genetella Way	West
31	31	1900-01-01 00:00:00.000	Angela Hernandez	angela.hernandez@baklaccustomer.org	18481880151	788 Aurora Avenue	Yamaguchi
32	32	1900-01-01 00:00:00.000	Andrea King	andrea.king@baklaccustomer.org	52525548078	1405 Ansonia Manor	Sierra Leone
33	33	1900-01-01 00:00:00.000	Brenda Wright	brenda.wright@baklaccustomer.org	74589494706	33 Genetella Way	Kanar
34	34	1900-01-01 00:00:00.000	Amy Lopez	amy.lopez@baklaccustomer.org	65770597774	178 Mandelgong Place	Utter Pradesh
35	35	1900-01-01 00:00:00.000	Amy Hall	amy.hall@baklaccustomer.org	61722222278	127 Turner Puma Manor	Nemete
36	36	1900-01-01 00:00:00.000	Rebecca Scott	rebecca.scott@baklaccustomer.org	70848338270	61 Tama Street	Osakima
37	37	1900-01-01 00:00:00.000	Virginia Green	virginia.green@baklaccustomer.org	44051215168	391 Calfee Drive	Mad-Pynes
38	38	1900-01-01 00:00:00.000	Kathleen Adams	kathleen.adams@baklaccustomer.org	48118274822	338 Manager Monaghan Lane	Maluku
39	39	1900-01-01 00:00:00.000	Pamela Baker	pamela.baker@baklaccustomer.org	91225720465	1480 Fukuyama Loop	Honan
40	40	1900-01-01 00:00:00.000	Martha Gonzalez	martha.gonzalez@baklaccustomer.org	48781838737	288 Cam Rana Parkway	Chiosia
41	41	1900-01-01 00:00:00.000	Debra Nelson	debra.nelson@baklaccustomer.org	37811883161	384 SantaAgusta Place	Episita Santo

**Gambar 3.2** Data Customer pada Tabel dim\_customer

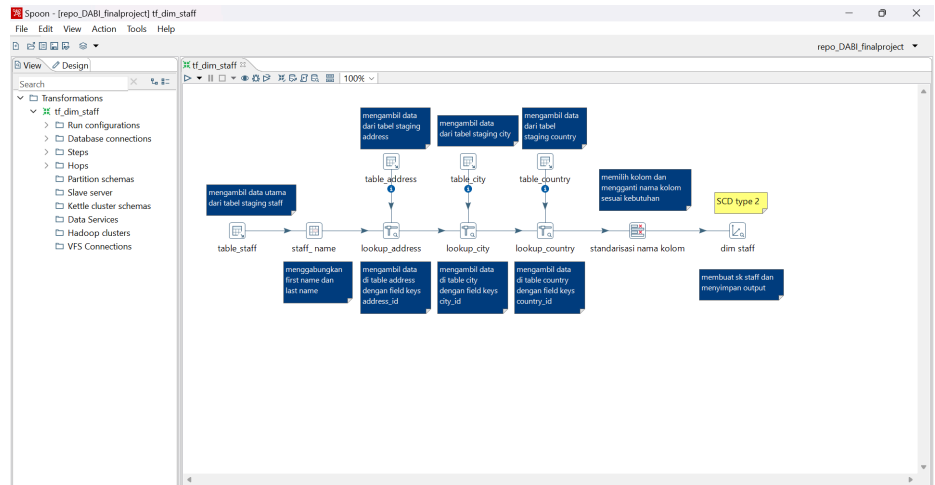
- Dimensi Film

- Bertujuan untuk melakukan analisis berdasarkan atribut film, seperti kategori, bahasa, dan aktor.
- Data diambil dari tabel staging seperti film, language, category, dan aktor.
- Transformasi meliputi penggabungan nama depan dan belakang aktor menjadi actor\_name serta perubahan tipe data atribut tertentu.
- Output disimpan di tabel dim\_film dengan penerapan SCD Type 2 untuk menyimpan perubahan data film.
- Gambar 3.3 dan 3.4 berikut merupakan tampilan dari transformasi tf\_dim\_film dan hasil data ditampilkan di tabel dim\_film.



Enter a SQL expression to filter results (use Ctrl+Space)												
	123 film	123 version	date from	date to	123 film_id	A2 title	A2 description	123 release_year	123 rental_year	123 length	A2 rating	A2 special features
1			1900-01-01 00:00:00	2199-12-31 23:59:59	1	A Royal Refinement of a Movie and a Hit	A Royal Refinement of a Movie and a Hit	2006	4.00	117	NC-17	[Trains]
2			1900-01-01 00:00:00	2199-12-31 23:59:59	133	Grocery Wonderful	A Royal Refinement of a Movie and a Hit	2006	4.00	179	NC-17	[Trains]
3			1900-01-01 00:00:00	2199-12-31 23:59:59	8	Alphabet Pollock	A Royal Refinement of a Movie and a Hit	2006	4.00	54	R	[Trains]
4			1900-01-01 00:00:00	2199-12-31 23:59:59	6	Eighty Encounters	A Royal Refinement of a Movie and a Hit	2006	4.00	73	PG-13	[Trains]
5			1900-01-01 00:00:00	2199-12-31 23:59:59	134	Adaptation Reflections	A Royal Refinement of a Movie and a Hit	2006	4.00	117	NC-17	[Trains]
6			1900-01-01 00:00:00	2199-12-31 23:59:59	4	A Royal Refinement of a Movie and a Hit	A Royal Refinement of a Movie and a Hit	2006	4.00	49	R	[Trains]
7			1900-01-01 00:00:00	2199-12-31 23:59:59	3	A Royal Refinement of a Movie and a Hit	A Royal Refinement of a Movie and a Hit	2006	4.00	49	R	[Trains]
8			1900-01-01 00:00:00	2199-12-31 23:59:59	3	Adaptation Reflections	A Royal Refinement of a Movie and a Hit	2006	2.00	180	PG	[Trains]
9			1900-01-01 00:00:00	2199-12-31 23:59:59	3	Adaptation Reflections	A Royal Refinement of a Movie and a Hit	2006	2.00	180	PG	[Trains]
10			1900-01-01 00:00:00	2199-12-31 23:59:59	6	Agresty Tournant	A Royal Refinement of a Movie and a Hit	2006	2.00	180	PG	[Trains]
11			1900-01-01 00:00:00	2199-12-31 23:59:59	6	Agresty Tournant	A Royal Refinement of a Movie and a Hit	2006	2.00	180	PG	[Trains]
12			1900-01-01 00:00:00	2199-12-31 23:59:59	9	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	2.00	114	NC-17	[Trains]
13			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
14			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
15			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
16			1900-01-01 00:00:00	2199-12-31 23:59:59	213	Delta Speed	A Royal Refinement of a Movie and a Hit	2006	0.00	158	R	[Trains]
17			1900-01-01 00:00:00	2199-12-31 23:59:59	213	Delta Speed	A Royal Refinement of a Movie and a Hit	2006	0.00	158	R	[Trains]
18			1900-01-01 00:00:00	2199-12-31 23:59:59	213	Delta Speed	A Royal Refinement of a Movie and a Hit	2006	0.00	158	R	[Trains]
19			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
20			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
21			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
22			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
23			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
24			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
25			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
26			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
27			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
28			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
29			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
30			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
31			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
32			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
33			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
34			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
35			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
36			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
37			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
38			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
39			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]
40			1900-01-01 00:00:00	2199-12-31 23:59:59	14	Alphabetic Calendar	A Royal Refinement of a Movie and a Hit	2006	4.00	68	PG-13	[Trains]

- **Dimensi Staff**
  - Digunakan untuk melacak kontribusi staf terhadap transaksi penyewaan.
  - Data diambil dari tabel staging staff, address, city, dan country.
  - Transformasi mencakup penggabungan nama depan dan belakang menjadi staff\_name serta penyesuaian nama kolom (contoh: email menjadi staff\_email).
  - Output disimpan di tabel dim\_staff.
  - Gambar 3.5 dan 3.6 berikut merupakan tampilan dari transformasi tf dim\_staff dan hasil data ditampung di tabel dim\_staff.

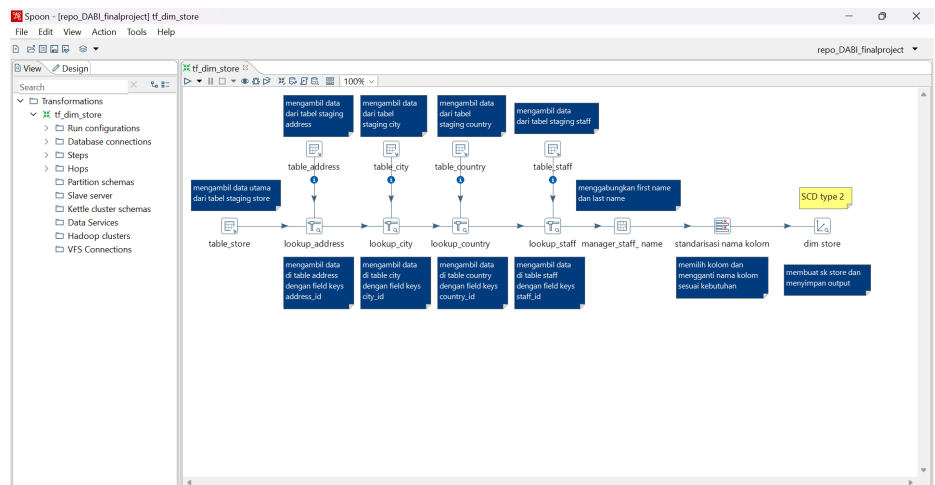


**Gambar 3.5** Tampilan Transformasi tf\_dim\_staff

sk_staff	version	date_from	date_to	staff_id	staff_name	staff_email	staff_address	staff_district	staff_city	staff_country	staff_postalcode	staff
1	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	1	Mike Hillyer	Mike.Hillyer@quibdatastaff.com	23 Wenhaven Lane	Alberta	Lethbridge	Canada	T1K0L1	1403333
2	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	2	Jon Stephens	Jon.Stephens@quibdatastaff.com	3411 1st Street Drive	CA D	Woodbridge	Australia	817225	877225

**Gambar 3.6** Data Staff pada Tabel dim\_staff

- Dimensi Store
  - Memungkinkan analisis performa berdasarkan lokasi toko.
  - Data diambil dari tabel staging store, address, city, dan staff.
  - Transformasi melibatkan lookup untuk manajer toko (manager\_staff\_name) dan penyesuaian atribut alamat toko.
  - Output disimpan di tabel dim\_store.
  - Gambar 3.7 dan 3.8 berikut merupakan tampilan dari transformasi tf\_dim\_store dan hasil data ditampung di tabel dim\_store.

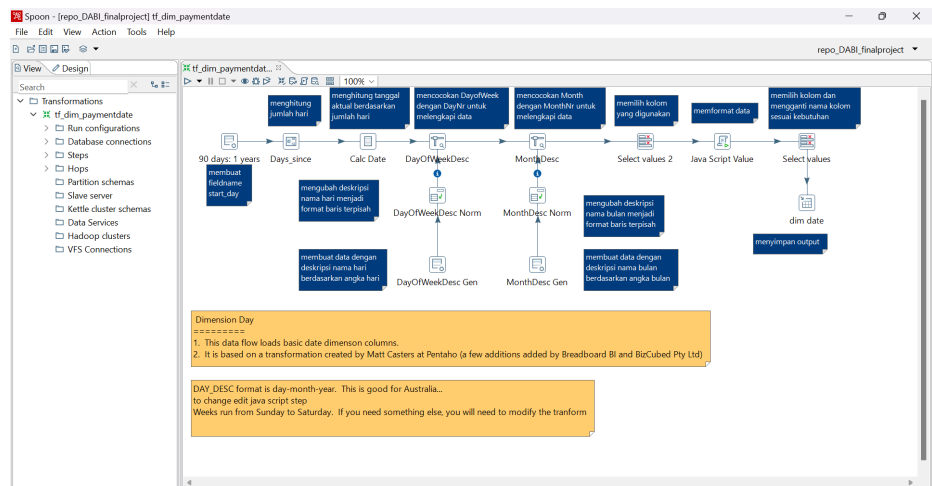


**Gambar 3.7** Tampilan Transformasi tf\_dim\_store

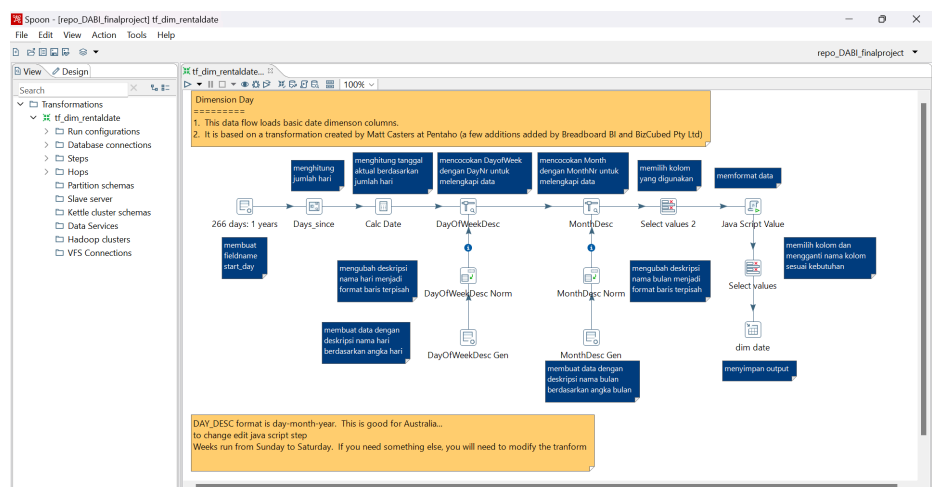
sk_store	version	date_from	date_to	store_id	manager_staff_name	store_address	store_district	store_city	store_country	store_postalcode
1	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	1	Mike Hillyer	47 MySakia Drive	Alberta	Lethbridge	Canada	T1K0L1
2	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	2	Jon Stephens	28 MySakia Boulevard	CA D	Woodbridge	Australia	817225

**Gambar 3.8** Data Store pada Tabel dim\_store

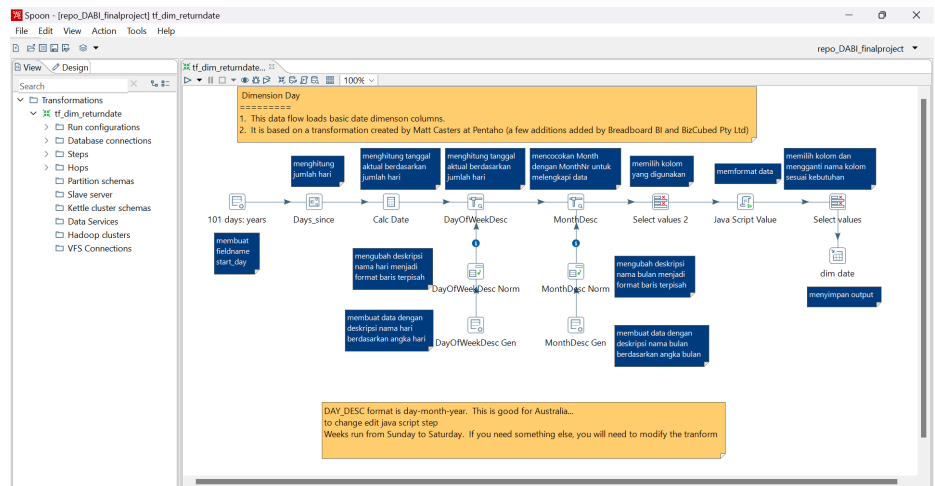
- Dimensi Waktu (Paymentdate, Rentaldate, Returndate)
  - Ketiga tabel dimensi waktu bertujuan untuk memberikan konteks temporal pada transaksi.
  - Data dihasilkan menggunakan transformasi di Pentaho, seperti generate rows untuk membuat tanggal, lalu dihitung deskripsi hari, bulan, dan kuartal.
  - Output disimpan di tabel dim\_paymentdate, dim\_rentaldate, dan dim\_returndate.
  - Gambar 3.9, 3.10, 3.11 berikut merupakan tampilan dari transformasi tf\_dim\_paymentdate, tf\_dim\_rentaldate, dan tf\_dim\_returndate.



**Gambar 3.9** Tampilan Transformasi tf\_dim\_paymentdate



**Gambar 3.10** Tampilan Transformasi tf\_dim\_rentaldate



**Gambar 3.11** Tampilan Transformasi tf\_dim\_returndate

- o Gambar 3.12, 3.13, 3.14 berikut merupakan tampilan dari hasil data ditampung di tabel dim\_paymentdate, dim\_rentaldate, dan dim\_returndate.

**Gambar 3.12** Data Film pada Tabel dim\_paymentdate

File: 2024-12-31 00:00:00 - Result: data.csv

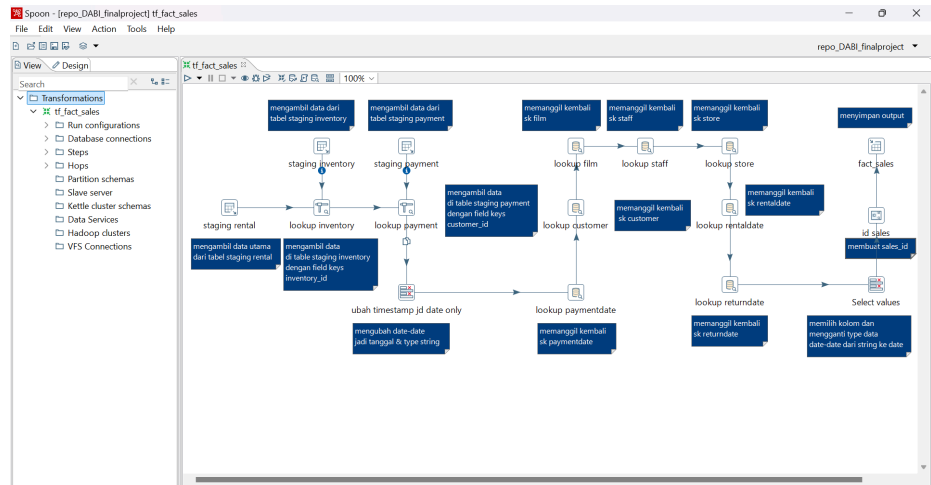
	16_remanensi	3_tanggal	45_deskopis_tanggal	133_16_tahun	133_tahun_angka	133_16_kuantal	133_kuantal_angka	45_kuantal	45_kuantal_tahun	133_16_kuantal	133_bulan_angka	45_bulan	133_16_bulan
1	20.050.520	2005-05-24 00:00:00.000	24-05-2005	2.005	2.005	2	2 02	02/2005	5	5	5	May	5
2	20.050.520	2005-05-25 00:00:00.000	25-05-2005	2.005	2.005	2	2 02	02/2005	5	5	5	May	5
3	20.050.520	2005-05-26 00:00:00.000	26-05-2005	2.005	2.005	2	2 02	02/2005	5	5	5	May	5
4	20.050.527	2005-05-27 00:00:00.000	27-05-2005	2.005	2.005	2	2 02	02/2005	5	5	5	May	5
5	20.050.520	2005-05-28 00:00:00.000	28-05-2005	2.005	2.005	2	2 02	02/2005	5	5	5	May	5
6	20.050.520	2005-05-29 00:00:00.000	29-05-2005	2.005	2.005	2	2 02	02/2005	5	5	5	May	5
7	20.050.520	2005-05-30 00:00:00.000	30-05-2005	2.005	2.005	2	2 02	02/2005	5	5	5	May	5
8	20.050.531	2005-05-31 00:00:00.000	31-05-2005	2.005	2.005	2	2 02	02/2005	5	5	5	May	5
9	20.050.601	2005-06-01 00:00:00.000	01-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
10	20.050.602	2005-06-02 00:00:00.000	02-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
11	20.050.601	2005-06-03 00:00:00.000	03-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
12	20.050.604	2005-06-04 00:00:00.000	04-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
13	20.050.603	2005-06-05 00:00:00.000	05-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
14	20.050.605	2005-06-06 00:00:00.000	06-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
15	20.050.607	2005-06-07 00:00:00.000	07-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
16	20.050.608	2005-06-08 00:00:00.000	08-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
17	20.050.609	2005-06-09 00:00:00.000	09-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
18	20.050.610	2005-06-10 00:00:00.000	10-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
19	20.050.611	2005-06-11 00:00:00.000	11-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
20	20.050.612	2005-06-12 00:00:00.000	12-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
21	20.050.611	2005-06-13 00:00:00.000	13-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
22	20.050.614	2005-06-14 00:00:00.000	14-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
23	20.050.614	2005-06-15 00:00:00.000	15-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
24	20.050.615	2005-06-16 00:00:00.000	16-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
25	20.050.617	2005-06-17 00:00:00.000	17-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
26	20.050.618	2005-06-18 00:00:00.000	18-06-2005	2.005	2.005	2	2 02	02/2005	6	6	6	June	6
27	20.050.619	2005-06-19 00:00:00.00											

### 3.3.3 Membuat Tabel Fakta

Tabel fakta adalah inti dari data warehouse, menyimpan data kuantitatif yang digunakan untuk analisis. Pada kasus ini, tabel fakta yang dibuat adalah:

- Berisi catatan transaksi penyewaan dan menghubungkan berbagai tabel dimensi melalui kolom foreign key.
- Data diambil dari tabel staging seperti rental, inventory, dan payment.
- Transformasi melibatkan:
  - Lookup ke tabel dimensi seperti dim\_customer, dim\_film, dan dim\_store.
  - Penggabungan atribut kuantitatif seperti amount dan penyesuaian tipe data tanggal menjadi format date.
  - Penambahan kolom sales\_id sebagai primary key menggunakan add sequence.
- Output disimpan di tabel fact\_sales.
- Gambar 3.15 dan 3.16 berikut merupakan tampilan dari transformasi fact\_sales dan hasil data ditampung di tabel fact\_sales.





**Gambar 3.15** Tampilan Transformasi tf\_fact\_sales

id	sk_paymentdate	sk_rentaldate	sk_returndate	sk_film	sk_customer	sk_staff	sk_store	amount	payment_date	rental_date	return_date	sales_id
1	2019-04-30	2019-04-30	2019-04-30	334	400	1	2	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	1
2	2019-04-30	2019-04-30	2019-04-30	337	334	2	1	2.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	2
3	2019-04-30	2019-04-30	2019-04-30	432	233	1	2	1.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	3
4	2019-04-30	2019-04-30	2019-04-30	615	549	1	1	0.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	4
5	2019-04-30	2019-04-30	2019-04-30	671	270	2	2	0	2017-05-14 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	5
6	2019-04-30	2019-04-30	2019-04-30	511	240	2	1	5.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	6
7	2019-04-30	2019-04-30	2019-04-30	566	127	1	1	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	7
8	2019-04-30	2019-04-30	2019-04-30	397	400	2	2	2.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	8
9	2019-04-30	2019-04-30	2019-04-30	971	142	2	1	0.99	2017-05-14 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	9
10	2019-04-30	2019-04-30	2019-04-30	348	262	2	1	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	10
11	2019-04-30	2019-04-30	2019-04-30	499	335	1	1	0.99	2017-05-14 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	11
12	2019-04-30	2019-04-30	2019-04-30	595	447	1	2	0.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	12
13	2019-04-30	2019-04-30	2019-04-30	671	335	1	1	0.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	13
14	2019-04-30	2019-04-30	2019-04-30	92	337	2	1	6.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	14
15	2019-04-30	2019-04-30	2019-04-30	166	375	1	2	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	15
16	2019-04-30	2019-04-30	2019-04-30	743	35	2	1	7.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	16
17	2019-04-30	2019-04-30	2019-04-30	423	427	1	2	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	17
18	2019-04-30	2019-04-30	2019-04-30	770	186	2	1	2.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	18
19	2019-04-30	2019-04-30	2019-04-30	35	388	2	1	0.99	2017-05-14 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	19
20	2019-04-30	2019-04-30	2019-04-30	162	192	2	1	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	20
21	2019-04-30	2019-04-30	2019-04-30	971	459	1	1	0.99	2017-05-14 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	21
22	2019-04-30	2019-04-30	2019-04-30	723	351	1	1	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	22
23	2019-04-30	2019-04-30	2019-04-30	864	38	2	2	5.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	23
24	2019-04-30	2019-04-30	2019-04-30	951	372	1	1	0.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	24
25	2019-04-30	2019-04-30	2019-04-30	272	362	2	2	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	25
26	2019-04-30	2019-04-30	2019-04-30	889	233	2	1	2.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	26
27	2019-04-30	2019-04-30	2019-04-30	156	45	2	2	4.99	2017-05-14 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	27
28	2019-04-30	2019-04-30	2019-04-30	820	431	1	1	2.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	28
29	2019-04-30	2019-04-30	2019-04-30	970	270	1	2	0.99	2017-05-14 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	29
30	2019-04-30	2019-04-30	2019-04-30	839	237	1	2	0.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	30
31	2019-04-30	2019-04-30	2019-04-30	368	273	1	1	10.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	31
32	2019-04-30	2019-04-30	2019-04-30	574	507	2	1	4.99	2017-05-14 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	32
33	2019-04-30	2019-04-30	2019-04-30	285	465	2	2	8.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	33
34	2019-04-30	2019-04-30	2019-04-30	289	89	1	2	2.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	34
35	2019-04-30	2019-04-30	2019-04-30	95	255	1	2	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	35
36	2019-04-30	2019-04-30	2019-04-30	560	303	1	1	1.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	36
37	2019-04-30	2019-04-30	2019-04-30	976	208	2	1	2.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	37
38	2019-04-30	2019-04-30	2019-04-30	770	444	1	1	0.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	38
39	2019-04-30	2019-04-30	2019-04-30	383	175	1	2	2.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	39
40	2019-04-30	2019-04-30	2019-04-30	90	524	2	1	6.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	40
41	2019-04-30	2019-04-30	2019-04-30	334	400	1	2	4.99	2017-04-30 00:00:00.000	2015-05-24 00:00:00.000	2015-05-24 00:00:00.000	41

**Gambar 3.16** Data Sales pada Tabel fact\_sales

### 3.3.4 Validasi dan Penyimpanan

Setelah transformasi selesai, semua tabel dimensi dan fakta disimpan di dalam schema dwh pada database dwh\_dvdrental. Hasilnya, data warehouse ini siap digunakan untuk tahap pembuatan datamart dan dashboard.

## 3.2 Ekstraksi Data ke MySQL

Untuk melakukan migrasi skema dwh dari database dwh\_dvdrental di DBeaver ke MySQL melalui XAMPP, pastikan langkah-langkah berikut dilakukan dengan benar. Pertama, pastikan layanan Apache dan MySQL di XAMPP Control Panel sudah dalam keadaan running. Selanjutnya, di DBeaver, lakukan backup schema dwh dari database dwh\_dvdrental ke dalam sebuah file SQL. Setelah itu, buka phpMyAdmin melalui localhost di peramban web dan buat database baru dengan nama "dwh". Setelah database berhasil dibuat, buka tab SQL pada database dwh di phpMyAdmin dan tempelkan seluruh query hasil backup schema dwh yang

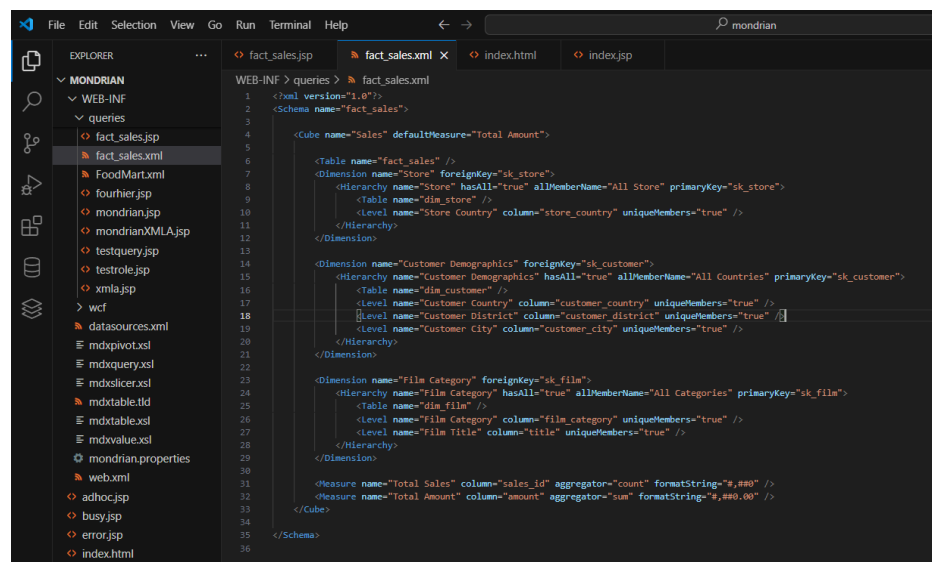
telah disimpan sebelumnya. Terakhir, jalankan query tersebut untuk memastikan semua data dan struktur dari skema dwh berhasil diimpor ke dalam database MySQL.

Proses ini dilakukan sebagai tahap awal sebelum melanjutkan ke langkah berikutnya, yaitu implementasi pembuatan CUBE dan OLAP menggunakan Mondrian Tomcat. Dengan struktur data yang sudah siap di database MySQL, proses data analysis dan penyajian informasi melalui Online Analytical Processing (OLAP) dapat dilakukan dengan lebih optimal menggunakan Mondrian yang dijalankan di Tomcat.

### 3.3 Implementasi CUBE dan OLAP

#### 3.3.1 Query

##### 1.) fact\_sales.xml



Gambar 3.17 fact\_sales.xml

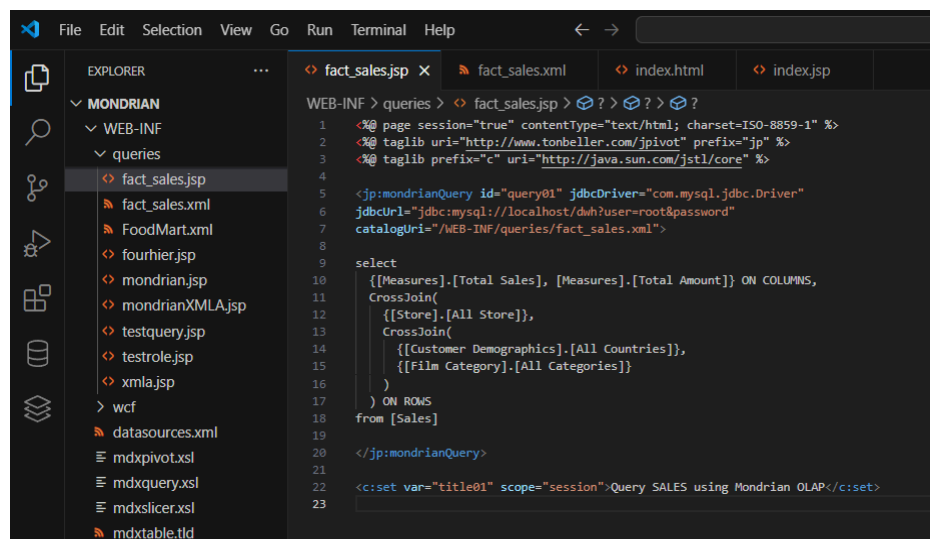
Gambar 3.17 di atas merupakan query dari fac\_sales.xml. File fact\_sales.xml bertindak sebagai konfigurasi skema untuk mendefinisikan struktur data cube yang akan dianalisis menggunakan Mondrian OLAP. Di dalam file ini, terdapat elemen Cube dengan nama Sales, yang menggunakan tabel fact\_sales sebagai sumber datanya. Cube ini memiliki dua measure utama, yaitu Total Sales, yang menghitung jumlah transaksi berdasarkan kolom sales\_id, dan Total Amount, yang menghitung total nilai penjualan berdasarkan kolom amount. Kedua measure ini menggunakan fungsi agregasi (count untuk Total Sales dan sum untuk Total Amount) serta format angka yang sudah ditentukan.

Selain itu, terdapat beberapa dimensi (dimensions) yang membantu mengorganisir data dalam bentuk hierarki. Pertama, Dimensi Store, yang

menggunakan kolom `sk_store` sebagai foreign key, memungkinkan analisis berdasarkan lokasi toko. Hierarki di dalamnya mencakup tingkatan seperti `store_country` untuk melihat data berdasarkan negara tempat toko berada. Kedua, Dimensi Customer Demographics, yang menggunakan kolom `sk_customer` sebagai foreign key, menyediakan informasi pelanggan yang lebih mendalam. Hierarkinya mencakup tingkatan negara pelanggan (`customer_country`), distrik pelanggan (`customer_district`), dan kota pelanggan (`customer_city`). Ketiga, Dimensi Film Category, dengan kolom `sk_film` sebagai foreign key, memungkinkan analisis berdasarkan kategori film (`film_category`) dan judul film (`title`).

Dengan struktur ini, setiap dimensi dan hierarki memungkinkan pengguna untuk mengeksplorasi data dalam berbagai sudut pandang, misalnya melihat total penjualan per negara, kota, atau berdasarkan kategori film tertentu. Konfigurasi dalam `fact_sales.xml` memastikan bahwa setiap level hierarki dan hubungan antar tabel didefinisikan dengan jelas sehingga query OLAP dapat dijalankan dengan efisien.

## 2.) `fact_sales.jsp`



```

1  <%@ page session="true" contentType="text/html; charset=ISO-8859-1" %>
2  <%@ taglib uri="http://www.tonbeller.com/jpivot" prefix="jp" %>
3  <%@ taglib prefix="c" uri="http://java.sun.com/jstl/core" %>
4
5  <jp:mondrianQuery id="query01" jdbcDriver="com.mysql.jdbc.Driver"
6  jdbcUrl="jdbc:mysql://localhost/dwh?user=root&password"
7  catalogUri="/WEB-INF/queries/fact_sales.xml">
8
9  select
10  [[Measures].[Total Sales], [Measures].[Total Amount]] ON COLUMNS,
11  CrossJoin(
12  [[Store].[All Store]],
13  CrossJoin(
14  [[Customer Demographics].[All Countries]],
15  [[Film Category].[All Categories]]
16  )
17  ) ON ROWS
18  from [Sales]
19
20 </jp:mondrianQuery>
21
22 <c:set var="title01" scope="session">Query SALES using Mondrian OLAP</c:set>
23

```

**Gambar 3.18** `fact_sales.jsp`

Gambar 3.18 di atas merupakan query dari `fact_sales.jsp`. File `fact_sales.jsp` bertindak sebagai antarmuka untuk mengeksekusi query MDX (Multidimensional Expressions) menggunakan konfigurasi skema yang sudah ditentukan dalam `fact_sales.xml`. Di awal file, terdapat konfigurasi koneksi menggunakan tag `jp:mondrianQuery`. Koneksi ini memanfaatkan JDBC (Java Database Connectivity) untuk menghubungkan aplikasi JSP dengan database yang digunakan Mondrian. String koneksi mengacu pada database lokal dengan

kredensial yang sesuai (jdbc:mysql://localhost/db), serta katalog yang merujuk pada file skema fact\_sales.xml.

Bagian penting dalam file ini adalah query MDX yang digunakan untuk mengambil data dari cube Sales. Query tersebut menampilkan dua measure di bagian kolom (ON COLUMNS), yaitu [Measures].[Total Sales] dan [Measures].[Total Amount], yang bertindak sebagai metrik utama dalam analisis. Sementara di bagian baris (ON ROWS), query menampilkan tiga dimensi utama: Store, Customer Demographics, dan Film Category. Setiap dimensi diwakili oleh hierarki tertingginya, seperti All Store, All Countries, dan All Categories.

Query MDX tersebut memungkinkan pengguna untuk melihat metrik penjualan berdasarkan kombinasi berbagai dimensi, misalnya total penjualan untuk setiap kategori film di negara tertentu atau distribusi jumlah transaksi berdasarkan lokasi toko. Hasil query ini kemudian diolah dan dapat ditampilkan dalam antarmuka web JSP.

### 3.3.2 Hasil

#### 1.) Tampilan Awal

##### Query SALES using Mondrian OLAP



			Measures	
Store	Customer Demographics	Film Category	Total Sales	Total Amount
All Store	All Countries	All Categories	16,044	59,398.30

Slicer:

[back to index](#)

**Gambar 3.19** Tampilan Awal Mondrian

#### 2.) Dimensi Store

##### Query SALES using Mondrian OLAP



			Measures	
Store	Customer Demographics	Film Category	Total Sales	Total Amount
All Store	All Countries	All Categories	16,044	59,398.30
Australia	All Countries	All Categories	8,121	30,112.05
Canada	All Countries	All Categories	7,923	29,286.25

Slicer:

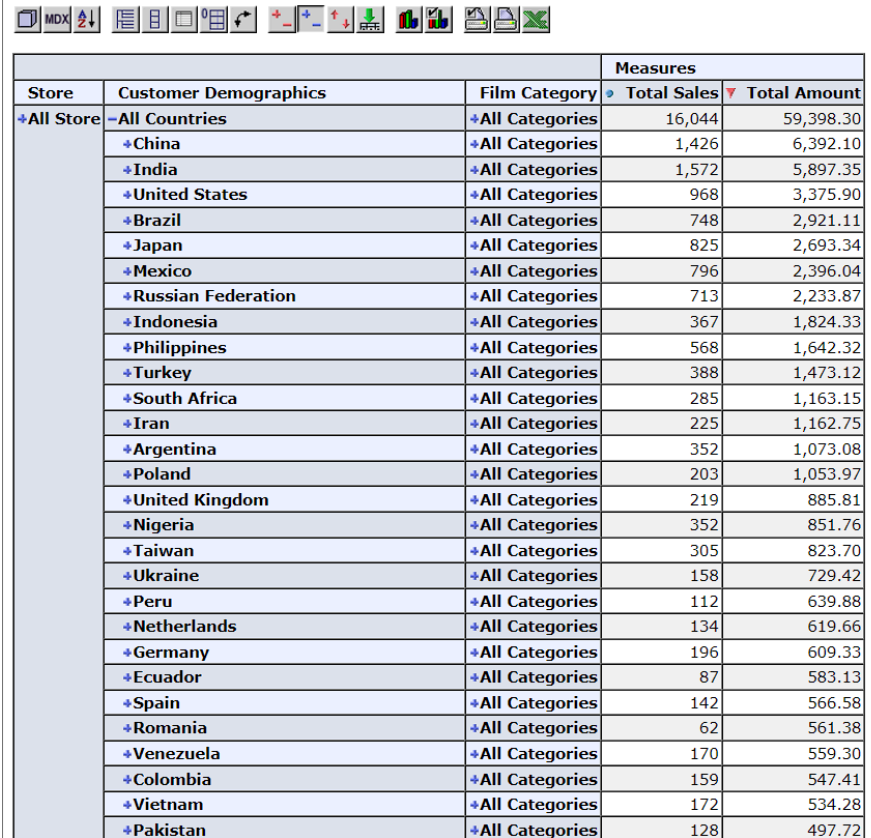
[back to index](#)

**Gambar 3.20** Tampilan Dimensi Store Mondrian

3.) Dimensi Customer Demographics

a. Customer Country

**Query SALES using Mondrian OLAP**



The screenshot shows the Mondrian OLAP interface with a toolbar at the top and a query result table below. The toolbar includes icons for MDX, filters, dimensions, measures, and other OLAP functions. The table displays data for the 'Store' dimension, specifically for 'All Store', across various 'Customer Demographics' (countries) and 'Film Category' (All Categories). The measures shown are 'Total Sales' and 'Total Amount'.

Store	Customer Demographics	Film Category	Measures
			◆ Total Sales ▼ Total Amount
◆ All Store	◆ All Countries	◆ All Categories	16,044 59,398.30
	◆ China	◆ All Categories	1,426 6,392.10
	◆ India	◆ All Categories	1,572 5,897.35
	◆ United States	◆ All Categories	968 3,375.90
	◆ Brazil	◆ All Categories	748 2,921.11
	◆ Japan	◆ All Categories	825 2,693.34
	◆ Mexico	◆ All Categories	796 2,396.04
	◆ Russian Federation	◆ All Categories	713 2,233.87
	◆ Indonesia	◆ All Categories	367 1,824.33
	◆ Philippines	◆ All Categories	568 1,642.32
	◆ Turkey	◆ All Categories	388 1,473.12
	◆ South Africa	◆ All Categories	285 1,163.15
	◆ Iran	◆ All Categories	225 1,162.75
	◆ Argentina	◆ All Categories	352 1,073.08
	◆ Poland	◆ All Categories	203 1,053.97
	◆ United Kingdom	◆ All Categories	219 885.81
	◆ Nigeria	◆ All Categories	352 851.76
	◆ Taiwan	◆ All Categories	305 823.70
	◆ Ukraine	◆ All Categories	158 729.42
	◆ Peru	◆ All Categories	112 639.88
	◆ Netherlands	◆ All Categories	134 619.66
	◆ Germany	◆ All Categories	196 609.33
	◆ Ecuador	◆ All Categories	87 583.13
	◆ Spain	◆ All Categories	142 566.58
	◆ Romania	◆ All Categories	62 561.38
	◆ Venezuela	◆ All Categories	170 559.30
	◆ Colombia	◆ All Categories	159 547.41
	◆ Vietnam	◆ All Categories	172 534.28
	◆ Pakistan	◆ All Categories	128 497.72

**Gambar 3.21** Tampilan Dimensi Customer Country Mondrian

b. Customer District

### Query SALES using Mondrian OLAP



			Measures	
Store	Customer Demographics	Film Category	Total Sales	Total Amount
+All Store	-All Countries	+All Categories	16,044	59,398.30
	-China	+All Categories	1,426	6,392.10
	+Shandong	+All Categories	236	1,171.64
	+Liaoning	+All Categories	82	618.18
	+Hubei	+All Categories	119	479.81
	+Jilin	+All Categories	63	430.37
	+Inner Mongolia	+All Categories	116	422.84
	+Sichuan	+All Categories	91	390.09
	+Henan	+All Categories	78	362.22
	+Guangdong	+All Categories	106	338.94
	+Zhejiang	+All Categories	53	284.47
	+Ningxia	+All Categories	30	269.70
	+Jiangsu	+All Categories	65	249.35
	+Shanxi	+All Categories	61	244.39
	+Hebei	+All Categories	32	223.68
	+Xinxiang	+All Categories	49	188.51
	+Heilongjiang	+All Categories	43	176.57
	+Hunan	+All Categories	58	175.78
	+Tianjin	+All Categories	23	114.77
	+Hainan	+All Categories	25	99.75
	+Fujian	+All Categories	14	69.86
	+Gansu	+All Categories	60	59.40
	+Jiangxi	+All Categories	22	21.78
	+India	+All Categories	1,572	5,897.35
	+United States	+All Categories	968	3,375.90
	+Brazil	+All Categories	748	2,921.11
	+Japan	+All Categories	825	2,693.34
	+Mexico	+All Categories	796	2,396.04

Gambar 3.22 Tampilan Dimensi Customer District Mondrian

#### c. Customer City

### Query SALES using Mondrian OLAP



			Measures	
Store	Customer Demographics	Film Category	Total Sales	Total Amount
+All Store	-All Countries	+All Categories	16,044	59,398.30
	-China	+All Categories	1,426	6,392.10
	-Shandong	+All Categories	236	1,171.64
	Yantai	+All Categories	20	219.80
	Liaocheng	+All Categories	34	203.66
	Xintai	+All Categories	24	167.76
	Weifang	+All Categories	29	144.71
	Junan	+All Categories	25	124.75
	Binzhou	+All Categories	33	98.67
	Laiwu	+All Categories	27	80.73
	Dongying	+All Categories	23	68.77
	Rizhao	+All Categories	21	62.79
	-Liaoning	+All Categories	82	618.18
	Jinzhou	+All Categories	30	329.70
	Tiefu	+All Categories	29	173.71
	Yingkou	+All Categories	23	114.77
	-Hubei	+All Categories	119	479.81
	Xiangfan	+All Categories	34	203.66
	Zaoyang	+All Categories	22	87.78
	Laohekou	+All Categories	26	77.74
	Ezhou	+All Categories	23	68.77
	Enshi	+All Categories	14	41.86

Gambar 3.23 Tampilan Dimensi Customer City Mondrian

#### 4.) Dimensi Film

a. Film Category

Query SALES using Mondrian OLAP



		Measures		
Store	Customer Demographics	Film Category	Total Sales	Total Amount
+All Store	+All Countries	-All Categories	16,044	59,398.30
		+Sports	1,179	4,472.66
		+Animation	1,166	4,299.80
		+Action	1,112	4,091.44
		+Sci-Fi	1,101	4,063.43
		+Family	1,096	4,030.49
		+Documentary	1,050	3,850.01
		+Drama	1,060	3,773.86
		+Foreign	1,033	3,679.12
		+Comedy	941	3,598.89
		+Children	945	3,572.89
		+Games	969	3,572.73
		+New	940	3,455.08
		+Classics	939	3,447.05
		+Music	830	3,202.95
		+Travel	837	3,151.02
		+Horror	846	3,136.88

Gambar 3.24 Tampilan Dimensi Film Category Mondrian

b. Film Tittle

Query SALES using Mondrian OLAP



		Measures		
Store	Customer Demographics	Film Category	Total Sales	Total Amount
+All Store	+All Countries	-All Categories	16,044	59,398.30
		-Sports	1,179	4,472.66
		Talented Homicide	29	123.73
		Philadelphia Wife	23	109.78
		Saturday Lambs	28	106.74
		Gleaming Jawbreaker	29	106.74
		Tights Dawn	26	98.75
		Chance Resurrection	27	98.74
		Aladdin Calendar	23	95.77
		Roses Treasure	28	93.73
		Shakespeare Saddle	22	91.79
		Northwest Polish	20	89.80
		Durham Panky	26	87.74
		Flatliners Killer	22	86.78
		Trading Pinocchio	25	86.75
		Tourist Pelican	22	84.79
		Congeniality Quest	22	82.78
		Seabiscuit Punk	26	81.75
		Dude Blindness	16	79.85
		Straight Hours	20	77.80
		Star Operation	21	76.80
		Pickup Driving	16	75.84
		Peak Forever	20	71.82
		Loser Hustler	15	70.85
		River Outlaw	20	70.82
		Novocaine Flight	19	70.82
		Mummy Creatures	16	69.85

Gambar 3.24 Tampilan Dimensi Film Tittle Mondrian

## **BAB IV**

### **PENUTUPAN**

#### **4.1 Kesimpulan**

Proyek pengembangan Data Warehouse dan OLAP untuk studi kasus DVD Rental berhasil diselesaikan dengan baik. Seluruh tahapan, mulai dari perancangan skema Data Warehouse dengan pendekatan Star Schema, proses ETL menggunakan Pentaho Data Integration, hingga implementasi OLAP Cube menggunakan Mondrian, telah dilakukan secara sistematis. Hasil akhir berupa cube OLAP memungkinkan analisis data multidimensi yang mendalam, memberikan wawasan strategis terkait performa toko, wilayah pelanggan, dan kategori film. Keseluruhan proyek ini memberikan pengalaman berharga dalam mengelola data secara efisien dan mendukung pengambilan keputusan berbasis data.

#### **4.2 Saran**

Untuk pengembangan lebih lanjut, beberapa saran yang dapat diberikan antara lain:

- 1.) Memperluas cakupan analisis dengan menambahkan dimensi atau measure baru, seperti analisis berdasarkan waktu transaksi secara lebih rinci.
- 2.) Mengintegrasikan alat visualisasi data lain seperti Tableau atau Power BI untuk meningkatkan kualitas visualisasi hasil analisis.
- 3.) Melakukan pengujian dengan dataset yang lebih besar untuk mengukur kinerja Data Warehouse dan OLAP dalam skala yang lebih luas.
- 4.) Menambahkan mekanisme otomatisasi untuk pembaruan data di Data Warehouse agar analisis dapat dilakukan secara real-time.



## **LAMPIRAN**

### **Lampiran 1: Link GitHub**

[https://github.com/hilmiarm/Konversi\\_DWO\\_22082010061\\_Hilmi-Arya-rafwa-Muhammad](https://github.com/hilmiarm/Konversi_DWO_22082010061_Hilmi-Arya-rafwa-Muhammad)