

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
Rafie Ahza Ghaisan_22082010192

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

DAFTAR ISI

DAFTAR ISI.....	1
BAB I : PENDAHULUAN.....	3
A Latar Belakang.....	3
B Rumusan Masalah.....	4
C Tujuan Kegiatan.....	4
D Batasan.....	4
BAB II : METODOLOGI.....	5
A. Daftar Pustaka.....	5
1. DVD Rental.....	5
2. Pentaho.....	5
3. XAMPP.....	5
4. MySQL.....	6
5. Visual Studio Code.....	6
6. Tomcat.....	6
7. Mondrian.....	7
B. Metode.....	7
1. Pembuatan Data Warehouse.....	7
2. Implementasi CUBE dan OLAP.....	7
BAB III : IMPLEMENTASI DAN HASIL.....	9
A. Pembuatan Data Warehouse.....	9
1. Tabel Dimensi Customer.....	9
Gambar 3.1 Transformasi tf_dim_customer.....	10
Gambar 3.2 Data Tabel dim_customer.....	10
2. Tabel Dimensi Film.....	10
Gambar 3.3 Transformasi tf_dim_film.....	11
Gambar 3.4 Data Tabel dim_film.....	11
3. Tabel Dimensi Staff.....	11
Gambar 3.5 Transformasi tf_dim_staff.....	11
Gambar 3.6 Data Tabel dim_staff.....	11
4. Tabel Dimensi Store.....	11
Gambar 3.7 Transformasi tf_dim_store.....	12
Gambar 3.8 Data Tabel dim_store.....	12
5. Tabel Dimensi Date.....	12
Gambar 3.9 Transformasi tf_dim_date.....	13
Gambar 3.10 Data Tabel dim_date.....	13

6. Tabel Fakta Rental.....	13
Gambar 3.11 Transformasi tf_fact_rental.....	14
Gambar 3.12 Data Tabel fact_rental.....	14
B. Implementasi CUBE dan OLAP.....	14
1. fact_rental.xml.....	14
Gambar 3.13 fact_rental.xml.....	15
2. fact_rental.jsp.....	15
Gambar 3.14 fact_rental.jsp.....	16
3. Output Mondrian.....	16
Gambar 3.15 Tampilan Awal Mondrian.....	16
Gambar 3.16 Tampilan Dimensi Store Mondrian.....	17
Gambar 3.17 Tampilan Dimensi Customer Country Mondrian.....	17
Gambar 3.18 Tampilan Dimensi Customer District Mondrian.....	18
Gambar 3.19 Tampilan Dimensi Customer City Mondrian.....	18
Gambar 3.20 Tampilan Dimensi Film Category Mondrian.....	18
BAB IV : PENUTUPAN.....	20
A. Kesimpulan.....	20
B. Saran.....	20
LAMPIRAN.....	22

BAB I : PENDAHULUAN

A Latar Belakang

Di tengah pesatnya perkembangan teknologi digital, manajemen data telah menjadi elemen kunci dalam mendukung pengambilan keputusan yang strategis dan berbasis bukti di dunia bisnis. Meskipun industri penyewaan DVD mungkin dianggap sebagai sektor yang sudah ketinggalan zaman di era layanan streaming modern, kenyataannya, sektor ini masih menyimpan potensi besar untuk dioptimalkan. Dengan pendekatan yang tepat dalam pengelolaan data, industri ini dapat meningkatkan efisiensi operasional, mengidentifikasi peluang pertumbuhan, serta memaksimalkan profitabilitas. Analisis data yang komprehensif memungkinkan perusahaan untuk memahami lebih dalam tentang preferensi pelanggan, tren penyewaan film, dan faktor-faktor yang mempengaruhi kinerja toko di berbagai lokasi.

Database DVD Rental menyediakan kumpulan data yang kaya dan terstruktur, mencakup informasi tentang pelanggan, transaksi penyewaan, kategori film, serta performa toko di berbagai wilayah. Data ini menjadi aset berharga yang, jika diolah dengan tepat, dapat menghasilkan wawasan berharga untuk mendukung pengambilan keputusan berbasis data. Dengan penerapan konsep Data Warehouse dan OLAP (Online Analytical Processing), perusahaan dapat menyusun data yang tersebar menjadi satu kesatuan yang terorganisir, mempermudah proses analisis untuk mengidentifikasi pola perilaku pelanggan, menganalisis tren penyewaan film dari waktu ke waktu, serta mengevaluasi kinerja toko berdasarkan parameter geografis dan kategori film.

Proyek ini bertujuan untuk merancang dan mengimplementasikan Data Warehouse yang andal dengan menggunakan alat-alat seperti PostgreSQL untuk manajemen basis data, DBeaver sebagai alat visualisasi dan eksplorasi database, serta Pentaho Data Integration (PDI) untuk proses ekstraksi, transformasi, dan pemuatan data (ETL). Selain itu, analisis multidimensi akan dilakukan menggunakan Mondrian OLAP, yang memungkinkan eksplorasi data dari berbagai sudut pandang dan tingkat agregasi yang berbeda.

Melalui integrasi teknologi ini, proyek ini diharapkan dapat memberikan wawasan mendalam yang mendukung pengambilan keputusan berbasis data. Wawasan tersebut mencakup identifikasi pelanggan dengan nilai tinggi, prediksi tren penyewaan di masa mendatang, serta optimalisasi strategi pemasaran dan inventaris film di setiap lokasi toko. Dengan pendekatan ini, industri penyewaan DVD dapat beradaptasi dengan lebih baik di tengah perubahan pasar dan terus bersaing di era digital.

B Rumusan Masalah

1. Bagaimana cara merancang struktur Data Warehouse yang optimal untuk menganalisis data dari studi kasus DVD Rental?
2. Bagaimana proses ETL (Extract, Transform, Load) dapat dijalankan secara efisien dengan menggunakan Pentaho Data Integration (PDI)?
3. Bagaimana penerapan OLAP (Mondrian OLAP) dapat digunakan untuk menganalisis performa penyewaan DVD berdasarkan kategori film, lokasi toko, dan profil pelanggan?
4. Bagaimana hasil analisis data dapat divisualisasikan dengan efektif untuk mendukung pengambilan keputusan yang strategis?

C Tujuan Kegiatan

1. Membangun skema Data Warehouse dengan pendekatan Star Schema, yang mencakup struktur tabel dimensi dan tabel fakta untuk mendukung analisis yang terfokus.
2. Mengoptimalkan proses ETL untuk mengekstrak, mentransformasi, dan memuat data dari sumber operasional ke dalam lingkungan Data Warehouse dengan efisien.
3. Merancang dan mengimplementasikan OLAP Cube untuk memungkinkan eksplorasi data secara multidimensi dan fleksibel.
4. Menyediakan wawasan yang mendalam dan dapat diandalkan untuk membantu pengambilan keputusan strategis dalam bisnis penyewaan DVD.

D Batasan

1. Studi ini hanya berfokus pada dataset dari DVD Rental.
2. Alat yang digunakan terbatas pada PostgreSQL, DBeaver, Pentaho Data Integration (PDI), Visual Studio Code, dan Mondrian OLAP.
3. Analisis data dibatasi pada tiga indikator utama: performa staff, distribusi geografis pelanggan, dan rating film.
4. Visualisasi data hanya akan dilakukan menggunakan antarmuka yang disediakan oleh Mondrian OLAP.

BAB II : METODOLOGI

A. Daftar Pustaka

1. DVD Rental

Studi kasus ini berfokus pada Analisis Penyewaan DVD, dengan tiga aspek utama yang menjadi perhatian: Performa Staff, Segmentasi Pelanggan, dan Kategori Film. Data yang digunakan berasal dari Database DVD Rental, yang sering dijadikan contoh dalam pembelajaran Sistem Manajemen Basis Data Relasional (RDBMS) serta analisis data di industri penyewaan DVD. Database ini mencakup berbagai tabel yang menyimpan informasi penting terkait pelanggan, stok film, transaksi penyewaan, pembayaran, dan kategori film.

DVD Rental sendiri berfungsi untuk menggambarkan seluruh operasional bisnis penyewaan DVD, mulai dari pengelolaan pelanggan, inventaris film, hingga transaksi harian. Beberapa tabel utama yang digunakan dalam analisis ini adalah Customer, Film, Rental, Inventory, Store, Payment, dan Category.

Fokus utama studi ini adalah untuk menggali pola penyewaan DVD yang bisa meningkatkan efisiensi operasional dan merancang strategi pemasaran yang lebih efektif. Analisis ini berusaha memahami bagaimana perilaku pelanggan dalam memilih film, bagaimana kinerja toko bervariasi antar lokasi, serta bagaimana kategori film memengaruhi tren penyewaan. Dengan wawasan yang didapat, diharapkan dapat mendukung pengambilan keputusan yang berbasis data dalam konteks bisnis penyewaan DVD.

2. Pentaho

Pentaho Data Integration (PDI), juga dikenal dengan nama Kettle, adalah alat open-source yang digunakan untuk ETL (Extract, Transform, Load). PDI memfasilitasi proses pengumpulan, pembersihan, dan transformasi data untuk dimuat ke dalam Data Warehouse atau sistem lainnya. Dengan antarmuka grafis berbasis drag-and-drop, pengguna dapat membangun pipeline data yang kompleks tanpa memerlukan keterampilan pemrograman mendalam. PDI mendukung integrasi dengan berbagai sumber data, termasuk basis data relasional, file CSV, dan layanan web. PDI sangat berguna dalam proyek-proyek data besar yang memerlukan alur data otomatis dan efisien, serta fleksibel dalam menggabungkan berbagai jenis sumber data.

3. XAMPP

XAMPP adalah paket perangkat lunak yang menyediakan Apache, MySQL, PHP, dan Perl dalam satu instalasi yang mudah digunakan. XAMPP dirancang untuk pengembangan lokal dan pengujian aplikasi web, memungkinkan pengembang untuk dengan cepat mengatur

server web lokal untuk pengembangan dan pengujian. Salah satu keunggulan XAMPP adalah kemudahan instalasi dan penggunaan, serta kemampuannya untuk menyediakan lingkungan yang lengkap untuk menjalankan aplikasi berbasis PHP dengan dukungan MySQL sebagai database backend. XAMPP sangat cocok untuk pengembangan aplikasi web secara lokal sebelum diterapkan di server produksi.

4. MySQL

MySQL adalah salah satu sistem RDBMS open-source yang paling populer dan banyak digunakan di dunia. MySQL menawarkan kecepatan, skalabilitas, dan kemudahan dalam pengelolaan basis data. Sebagai bagian dari Oracle Corporation, MySQL mendukung berbagai fitur standar SQL, termasuk transaksi, penguncian baris, dan replikasi basis data. MySQL digunakan dalam banyak aplikasi web, seringkali sebagai bagian dari stack LAMP (Linux, Apache, MySQL, PHP/Python/Perl). MySQL juga mendukung berbagai jenis penyimpanan data dan dapat diintegrasikan dengan berbagai aplikasi pengelolaan data.

5. Visual Studio Code

Visual Studio Code (VS Code) adalah editor teks sumber terbuka yang ringan, namun sangat kuat, yang digunakan untuk pengembangan perangkat lunak. Dikenal karena kemampuannya untuk mendukung berbagai bahasa pemrograman melalui ekstensi, termasuk JavaScript, Python, SQL, dan banyak lagi. VS Code menawarkan berbagai fitur seperti debugging, kontrol versi Git, dan integrasi terminal yang membuatnya menjadi pilihan utama di kalangan pengembang perangkat lunak. Selain itu, VS Code memiliki banyak ekstensi yang memungkinkan integrasi dengan alat lain seperti Docker, Kubernetes, dan alat pengembangan lainnya.

6. Tomcat

Tomcat adalah server aplikasi open-source yang dikembangkan oleh Apache Software Foundation dan dirancang untuk menjalankan aplikasi Java Servlet dan JavaServer Pages (JSP). Tomcat berfungsi sebagai server HTTP dan aplikasinya sering digunakan untuk meng-host aplikasi berbasis Java di web. Dengan performa tinggi dan kemudahan penggunaan, Tomcat menjadi pilihan utama bagi pengembang yang membangun aplikasi web berbasis Java. Tomcat mendukung berbagai fitur keamanan dan manajemen sesi, serta kemampuan untuk diintegrasikan dengan alat manajemen aplikasi dan basis data lainnya. Tomcat juga dapat diintegrasikan dengan JDBC untuk menghubungkan aplikasi dengan RDBMS.

7. Mondrian

Mondrian adalah server OLAP (Online Analytical Processing) open-source yang memungkinkan analisis data multidimensi dalam konteks Data Warehousing. Mondrian memungkinkan pengguna untuk melakukan analisis data yang kompleks melalui struktur dimensi yang dapat disesuaikan. Mondrian digunakan untuk membuat OLAP Cubes yang memungkinkan pemrosesan data dalam beberapa dimensi untuk memudahkan analisis. Dengan menggunakan bahasa query MDX (MultiDimensional eXpressions), Mondrian memungkinkan pengambilan keputusan berbasis data yang lebih dalam dan analisis data yang lebih canggih. Mondrian sering digunakan dalam aplikasi Business Intelligence (BI) untuk menganalisis data yang berasal dari berbagai sumber.

B. Metode

1. Pembuatan Data Warehouse

Tahap perancangan Data Warehouse berfokus pada pengorganisasian dan penyusunan data dalam struktur yang lebih efisien dan siap untuk analisis. Proses dimulai dengan pembuatan database `dwh_dvdrental` menggunakan DBeaver, diikuti dengan pembuatan skema `dwh` di dalam database tersebut. Skema ini bertindak sebagai kerangka kerja utama yang menyimpan tabel dimensi dan tabel fakta. Tabel dimensi yang dikembangkan meliputi `customers`, `film`, `staff`, `store`, serta tabel yang berhubungan dengan waktu yaitu `date`. Masing-masing tabel dimensi dirancang untuk menyimpan atribut yang menggambarkan entitas terkait, sementara tabel fakta seperti `rental` menyimpan data transaksi yang menjadi fokus utama analisis. Selama proses transformasi data, langkah-langkah dilakukan dengan hati-hati untuk menjaga konsistensi dan akurasi data.

2. Implementasi CUBE dan OLAP

Proses implementasi OLAP Cube dimulai dengan pembuatan file konfigurasi `fact_rental.xml`, yang berfungsi untuk mendefinisikan struktur cube OLAP. Di dalam file ini, berbagai dimensi seperti `Store`, `Customer Demographics`, dan `Film Category` diatur, bersama dengan measure seperti `Total Sales` dan `Total Amount`, yang akan digunakan untuk analisis data. Kemudian, dibuatlah file `fact_rental.jsp` untuk menghubungkan cube dengan database MySQL melalui JDBC. File ini juga menyertakan query MDX yang digunakan untuk mengambil data dari cube OLAP dan menampilkannya dalam bentuk yang dapat diakses melalui antarmuka web.

Setelah selesai, kedua file tersebut disalin ke dalam direktori `webapps` di server Tomcat. Server Tomcat dijalankan menggunakan XAMPP, dan antarmuka OLAP dapat diakses melalui browser untuk memastikan bahwa cube telah diatur dengan benar. Langkah terakhir adalah

menjalankan serangkaian pengujian untuk memastikan bahwa dimensi, measure, dan query MDX berfungsi dengan tepat sesuai harapan.

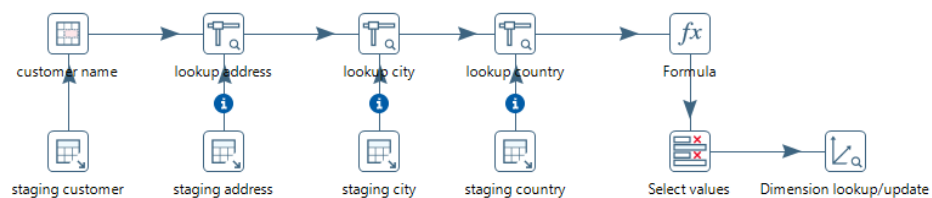
BAB III : IMPLEMENTASI DAN HASIL

A. Pembuatan 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. Langkah awal dimulai dengan pembuatan database baru bernama `dwh_dvdrental` menggunakan DBEaver, diikuti dengan pembuatan skema `dwh` di dalam database tersebut, yang berfungsi sebagai kerangka utama untuk menyimpan tabel dimensi dan tabel fakta.

1. Tabel Dimensi Customer

Tabel dimensi dirancang untuk menyimpan atribut deskriptif yang memberikan konteks pada data. Dimensi pelanggan (`dim_customer`) dibuat untuk memahami dan mengelompokkan data pelanggan berdasarkan atribut seperti nama, alamat, dan status aktif. Data ini diambil dari tabel staging seperti `customer`, `address`, `city`, dan `country`, yang kemudian melalui proses transformasi untuk menggabungkan nama depan dan belakang menjadi `customer_name`, serta penyesuaian nama kolom demi konsistensi. Hasil transformasi disimpan dalam tabel `dim_customer` dengan penerapan SCD Type 2 untuk menangani perubahan historis data pelanggan. Gambar 3.1 menunjukkan tampilan transformasi `tf_dim_customer`, sedangkan Gambar 3.2 menampilkan hasil data dalam tabel `dim_customer`.



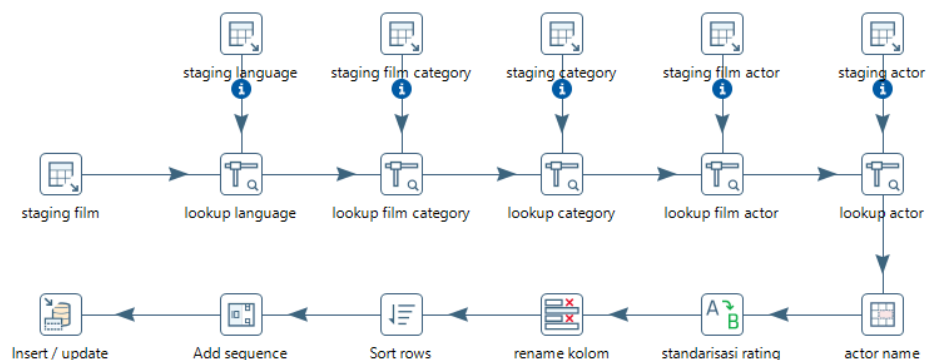
Gambar 3.1 Transformasi tf_dim_customer

id_customer	date_from	date_to	customer_id	customer_name	customer_email	customer_phone	customer_address	customer_district	customer_city
1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	1	Irene Ely	irene.ely@khalidcustomer.org	3833311987	1001 Gombakgates Street	10011	Purnani
2	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	2	Mary Smith	mary.smith@khalidcustomer.org	2830384290	1913 Hanco Way	Nagasaki	Sasab
3	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	3	Patricia Johnson	patricia.johnson@khalidcustomer.org	83803538668	1121 Luga Avenue	California	San Be
4	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	4	Linda Williams	linda.williams@khalidcustomer.org	44847719049	692 Joliet Street	Athiba	Athen
5	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	5	Barbara Jones	barbara.jones@khalidcustomer.org	70581480327	1566 Inegi Manor	Mandalay	Mying
6	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	6	Elizabeth Brown	elizabeth.brown@khalidcustomer.org	18655648674	53 Jofu Parkway	Nanturo	Nantur
7	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	7	Jennifer Davis	jennifer.davis@khalidcustomer.org	98652828454	1793 Lantago de Compostela Way	Texas	Lando
8	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	8	Maria Miller	maria.miller@khalidcustomer.org	71857120373	900 Santiago de Compostela Parkway	Central Serbia	Kraguj
9	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	9	Susan Wilson	susan.wilson@khalidcustomer.org	657382285970	478 Joliet Way	Hamilton	Hamth
10	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	10	Margaret Moore	margaret.moore@khalidcustomer.org	38625723569	413 Groove Drive	Musquet	Maysa
11	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	11	Dorothy Taylor	dorothy.taylor@khalidcustomer.org	64856898185	1531 Sol Drive	Esfahan	Esfaha
12	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	12	Lisa Anderson	lisa.anderson@khalidcustomer.org	63529727745	1542 Terlec Parkway	Kanagawa	Sagam
13	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	13	Nancy Thomas	nancy.thomas@khalidcustomer.org	46269801074	888 Bhopal Manor	Haryana	Vimpor
14	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	14	Karen Jackson	karen.jackson@khalidcustomer.org	69947967938	279 Amosby Parkway	Omanaya	Oman
15	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	15	Betty White	betty.white@khalidcustomer.org	51733814235	770 Bydgetz Avenue	California	Citisi
16	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	16	Helen Harris	helen.harris@khalidcustomer.org	99081110754	419 Igan Lane	Madhya Pradesh	Bhopa
17	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	17	Smitha Martin	smitha.martin@khalidcustomer.org	94813213387	380 Toulouse Parkway	England	Southa
18	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	18	Donna Thompson	donna.thompson@khalidcustomer.org	40772414682	270 Toulon Boulevard	Kalmykia	Eliota
19	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	19	Carol Garcia	carol.garcia@khalidcustomer.org	74791594869	330 East Avenue	Kaduna	Kadun
20	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	20	Ruth Martinez	ruth.martinez@khalidcustomer.org	27327217861	1817 Lim-Letter Avenue	Northwest Cape	Toriba
21	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	21	Sharon Robinson	sharon.robinson@khalidcustomer.org	14445889182	1688 Okawa Way	Northwest Border Prov	Marda
22	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	22	Michelle Clark	michelle.clark@khalidcustomer.org	88277379083	262 A Conna (La Conna) Parkway	Dhaka	Tangri
23	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	23	Leona Rodriguez	leona.rodriguez@khalidcustomer.org	16186821422	21 Charlotte Avenue Street	Kabul-Sal-Zeminton-2	Sel
24	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	24	Sarah Lewis	sarah.lewis@khalidcustomer.org	90273129323	1780 Biron Boulevard	Liepaja	Liepaja
25	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	25	Kimberly Lee	kimberly.lee@khalidcustomer.org	95472018745	96 Taluna Way	Cordoba	Cordib
26	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	26	Deborah Walker	deborah.walker@khalidcustomer.org	19499194769	834 San Felipe de Puerto Plata Street	Sind	Shiray
27	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	27	Jessica Hall	jessica.hall@khalidcustomer.org	98800977962	18 Duisburg Boulevard	Cile de	Cile de
28	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	28	Shirley Allen	shirley.allen@khalidcustomer.org	68336675023	217 Bethabeta Place	Southern Mindanao	Davai
29	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	29	Cynthia Young	cynthia.young@khalidcustomer.org	87822882805	1425 Siddigapal Manor	Bihar	Shung
30	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	30	Angela Hernandez	angela.hernandez@khalidcustomer.org	18481880151	786 Aurora Avenue	Yamaguchi	Shung
31	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	31	Melissa King	melissa.king@khalidcustomer.org	52525540978	1668 Angolis Street	Taipei	Lungth
32	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	32	Brenda Wright	brenda.wright@khalidcustomer.org	74299948749	55 Gorontalo Way	West Borneo	Kamir
33	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	33	Anna Lopez	anna.lopez@khalidcustomer.org	62769891751	179 Mandakung Place	Utara Pradesh	Purnai
34	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	34	Anna Hill	anna.hill@khalidcustomer.org	91182720578	127 Purnas (Purnas) Manor	Piemonte	Aleosa
35	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	35	Rebecca Scott	rebecca.scott@khalidcustomer.org	70440318370	81 Terna Street	Dharyana	Kurath
36	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	36	Virginia Green	virginia.green@khalidcustomer.org	48512131868	391 Collis Drive	Mid-Jyrna	Toulo
37	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	37	Kathleen Adams	kathleen.adams@khalidcustomer.org	48118527862	334 Munger (Munghy) Lane	Markazi	Akak
38	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	38	Pamela Baker	pamela.baker@khalidcustomer.org	91225758453	140 Fukuyama Loop	Hunan	Nanya
39	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	39	Maria Gonzalez	maria.gonzalez@khalidcustomer.org	48076302973	289 Cam East Parkway	Chiosini	Cosim

Gambar 3.2 Data Tabel dim_customer

2. Tabel Dimensi Film

Dimensi lainnya, seperti dim_film, dirancang untuk analisis berdasarkan atribut film, termasuk kategori, bahasa, dan aktor. Data ini berasal dari tabel staging seperti film, language, category, dan actor, yang ditransformasi dengan penggabungan nama aktor menjadi actor_name serta pengubahan tipe data atribut tertentu. Hasil akhirnya disimpan dalam tabel dim_film. Gambar 3.3 menunjukkan tampilan transformasi tf_dim_film, dan Gambar 3.4 menampilkan hasil data dalam tabel dim_film.



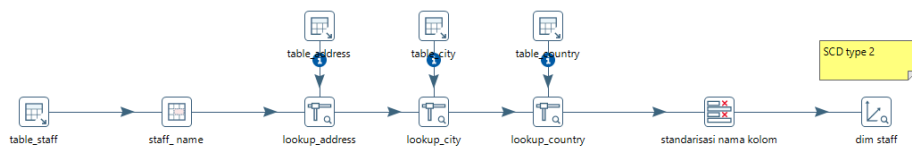
Gambar 3.3 Transformasi tf_dim_film

sk_film	version	date_from	date_to	id_film	title	description	release_year	rental_rate	length	rating	special_features
1	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	133	Chamber Italian	A Faithful Reflection of a Moose And a Hus	2006	4.99	117	NC-17	(Trailers)
2	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	204	Groove Wonderful	A Epic Drama of a Cat And a Explorer who	2006	4.99	49	R	(Behind the Scenes)
3	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	8	Airport Pollock	A Epic Tale of a Moose And a Girl who mu	2006	4.99	54	R	(Trailers)
4	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	96	Bright Encounters	A Faithful View of a Wilderness And a Fern	2006	4.99	71	PG-13	(Trailers)
5	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	1	Academy Discour	A Epic Drama of a Feminist And a Mad Sci	2006	0.99	86	PG	(Deleted Scenes), Behind the Sc
6	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	2	Acc Goldfinger	A Astounding Epistle of a Database Admin	2006	4.99	48	G	(Trailers, Deleted Scenes)
7	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	3	Adaptation Helic	A Astounding Reflection of a Lumberjack	2006	3.99	90	NC-17	(Trailers, Deleted Scenes)
8	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	4	After Prejudice	A Fearful Documentary of a Frisbee And a	2006	2.99	117	G	(Commentaries, Behind the Scen
9	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	5	African Egg	A Fast-Paced Documentary of a Peatty Che	2006	2.99	130	G	(Deleted Scenes)
10	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	6	Agent Trueman	A Intrepid Panorama of a Robot And a Boy	2006	2.99	109	PG	(Deleted Scenes)
11	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	7	Amplified Sierra	A Touching Saga of a Hunter And a Butler	2006	4.99	62	PG-13	(Trailers, Deleted Scenes)
12	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	9	Alabama Devil	A Thoughtful Panorama of a Database Adm	2006	2.99	114	PG-13	(Trailers, Deleted Scenes)
13	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	10	Aladdin Celestee	A Action-Packed Tale of a Man And a Lur	2006	4.99	82	NC-17	(Trailers, Deleted Scenes)
14	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	11	Alone Videotape	A Boring Epistle of a Butler And a Cat who	2006	0.99	126	G	(Commentaries, Behind the Scen
15	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	12	Alaska Phantom	A Fearful Saga of a Hunter And a Peatty C	2006	0.99	136	PG	(Commentaries, Deleted Scenes)
16	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	213	Date Speed	A Touching Saga of a Composer And a Mo	2006	0.99	104	R	(Commentaries)
17	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	13	Ali Forster	A Action-Packed Drama of a Dentist And a	2006	4.99	150	PG	(Deleted Scenes), Behind the Sc
18	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	14	Alice Fantasia	A Emotional Drama of a Shark And a Dal	2006	0.99	94	NC-17	(Trailers, Deleted Scenes), Behind
19	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	15	Allen Center	A Brilliant Drama of a Cat And a Mad Sci	2006	2.99	46	NC-17	(Trailers, Commentaries, Behind t
20	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	16	Alley Location	A Fast-Paced Drama of a Robot And a Car	2006	2.99	180	NC-17	(Trailers, Commentaries)
21	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	17	Alone Trip	A Fast-Paced Character Study of a Compo	2006	0.99	62	R	(Trailers, Behind the Scenes)
22	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	18	Alter Victory	A Thoughtful Drama of a Composer And a	2006	0.99	57	PG-13	(Trailers, Behind the Scenes)
23	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	19	Amateur Holy	A Emotional Display of a Pioneer And a Pa	2006	0.99	113	PG	(Commentaries, Deleted Scenes)
24	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	20	Amelie Hellfighters	A Boring Drama of a Woman And a Squa	2006	4.99	79	R	(Commentaries, Deleted Scenes)
25	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	21	American Circus	A Insightful Drama of a Girl And a Adonis	2006	4.99	129	R	(Commentaries, Behind the Scen
26	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	22	Amateur Madwoman	A Emotional Character Study of a Dentist	2006	2.99	83	G	(Commentaries, Behind the Scen
27	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	23	Aracoda Confessions	A Lachrymose Display of a Dentist And a D	2006	0.99	62	R	(Trailers, Deleted Scenes)
28	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	24	Analyze Housiers	A Thoughtful Display of a Explorer And a F	2006	2.99	181	R	(Trailers, Behind the Scenes)
29	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	25	Angels Life	A Thoughtful Display of a Woman And a F	2006	2.99	74	G	(Trailers)
30	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	26	Annie Identity	A Amazing Panorama of a Peatty Chef And	2006	0.99	68	G	(Commentaries, Deleted Scenes)
31	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	27	Anonymous Human	A Amazing Reflection of a Database Admin	2006	0.99	179	NC-17	(Deleted Scenes), Behind the Sc
32	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	28	Antihem Lake	A Touching Panorama of a Wilderness And a	2006	4.99	11	PG-13	(Deleted Scenes), Behind the Sc
33	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	29	Antihem Tomatoes	A Faithful View of a Wilderness And a Fern	2006	2.99	168	NC-17	(Trailers, Commentaries, Deleted)
34	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	30	Anything Savannah	A Epic Story of a Peatty Chef And a Woma	2006	2.99	62	R	(Trailers, Deleted Scenes), Behind
35	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	31	Apache Drive	A Awesome Inspiring Reflection of a Peatty Che	2006	4.99	152	NC-17	(Commentaries, Deleted Scenes)
36	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	32	Apocalypse Trainings	A Astounding Story of a Dog And a Squa	2006	4.99	118	R	(Trailers, Commentaries)
37	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	33	Apullo Team	A Action-Packed Reflection of a Crusade	2006	2.99	153	PG-13	(Trailers, Commentaries, Deleted)
38	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	34	Arabia Dogma	A Touching Epistle of a Madman And a Ma	2006	0.99	62	NC-17	(Commentaries, Deleted Scenes)
39	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	35	Awakened Robbery	A Action-Packed Reflection of a Peatty Che	2006	0.99	147	PG-13	(Trailers, Deleted Scenes), Behind

Gambar 3.4 Data Tabel dim_film

3. Tabel Dimensi Staff

Dimensi staf (dim_staff) berfokus pada pelacakan kontribusi staf terhadap transaksi penyewaan. Data dari tabel staging seperti staff, address, city, dan country ditransformasi untuk menggabungkan nama depan dan belakang menjadi staff_name. Hasilnya dapat dilihat pada Gambar 3.5 untuk tampilan transformasi tf_dim_staff, dan Gambar 3.6 untuk hasil data dalam tabel dim_staff.



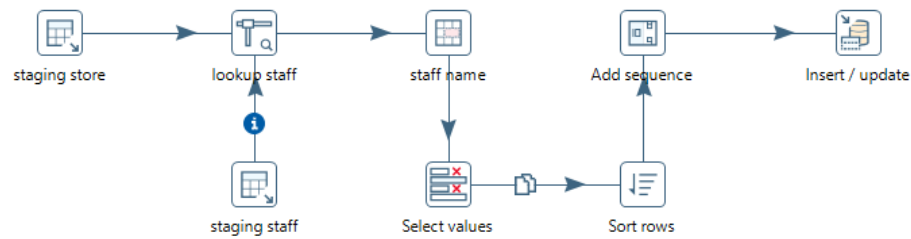
Gambar 3.5 Transformasi tf_dim_staff

sk_staff	version	date_from	date_to	id_staff	staff_name	staff_email	staff_address	staff_district	staff_city	staff_country
1	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	1	Mike Hilmyer	Mike.Hilmyer@sakilastaff.com	23 Workhaven Lane	Alberta	Lethbridge	Canada
2	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	2	Jon Stephens	Jon.Stephens@sakilastaff.com	1411 Lillydale Drive	QLD	Woodridge	Australia

Gambar 3.6 Data Tabel dim_staff

4. Tabel Dimensi Store

Dimensi toko (dim_store) memungkinkan analisis performa berdasarkan lokasi toko. Transformasi data untuk dimensi ini melibatkan proses lookup untuk menambahkan atribut seperti manager_staff_name dan penyesuaian atribut alamat toko, dengan hasil akhir disimpan di tabel dim_store. Gambar 3.7 menunjukkan tampilan transformasi tf_dim_store, sementara Gambar 3.8 menampilkan hasil data dalam tabel dim_store.



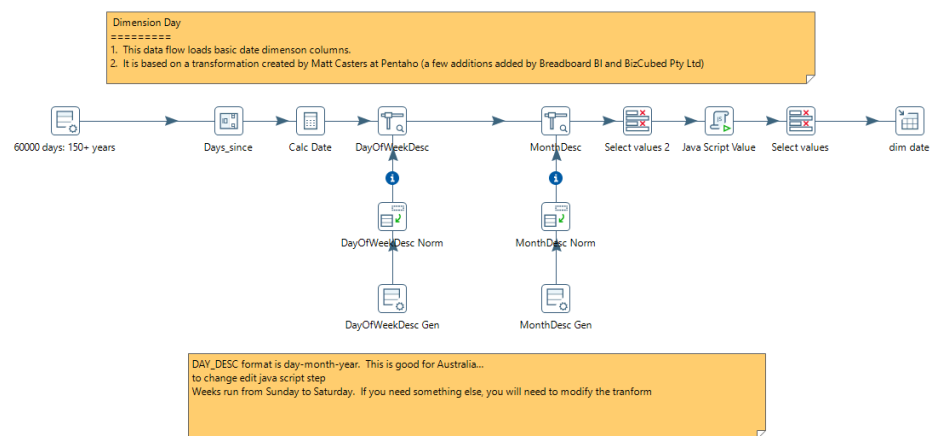
Gambar 3.7 Transformasi tf_dim_store

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 MySakila Drive	Alberta	Lethbridge	Canada	[NULL]
2	1	1900-01-01 00:00:00.000	2199-12-31 23:59:59.999	2	Jon Stephens	28 MySQL Boulevard	QLD	Woodridge	Australia	[NULL]

Gambar 3.8 Data Tabel dim_store

5. Tabel Dimensi Date

Tabel dimensi waktu (dim_date) memberikan konteks temporal pada transaksi. Data untuk tabel ini dihasilkan menggunakan transformasi Pentaho, seperti generate rows untuk membuat rentang tanggal, yang kemudian dihitung deskripsi harian, bulanan, dan kuartalannya. Gambar 3.9 menunjukkan tampilan transformasi tf_dim_date, sedangkan Gambar 3.10 menampilkan hasil data dalam tabel dim_date.



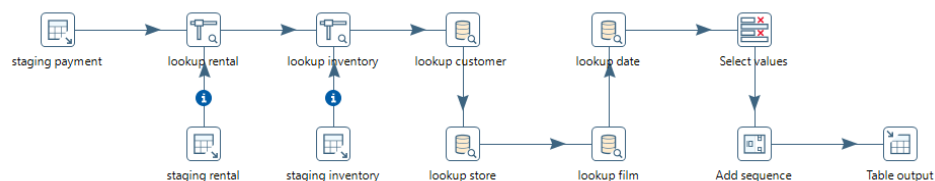
Gambar 3.9 Transformasi tf_dim_date

id	tf_dim_date	tanggal	deskripsi_tanggal	th_tahun	th_kuartal	kuartal_angka	kuartal_tahun	th_bulan	bulan_angka	bulan_min
1	20190101	2019-01-01 00:00:00.000	01-01-2019	2,005	2,005	2	2 Q1	012019	1	1 Jan
2	20190102	2019-01-02 00:00:00.000	02-01-2019	2,005	2,005	2	2 Q1	022019	2	2 Jan
3	20190103	2019-01-03 00:00:00.000	03-01-2019	2,005	2,005	2	2 Q1	032019	3	3 Jan
4	20190104	2019-01-04 00:00:00.000	04-01-2019	2,005	2,005	2	2 Q1	042019	4	4 Jan
5	20190105	2019-01-05 00:00:00.000	05-01-2019	2,005	2,005	2	2 Q1	052019	5	5 Jan
6	20190106	2019-01-06 00:00:00.000	06-01-2019	2,005	2,005	2	2 Q1	062019	6	6 Jan
7	20190107	2019-01-07 00:00:00.000	07-01-2019	2,005	2,005	2	2 Q1	072019	7	7 Jan
8	20190108	2019-01-08 00:00:00.000	08-01-2019	2,005	2,005	2	2 Q1	082019	8	8 Jan
9	20190109	2019-01-09 00:00:00.000	09-01-2019	2,005	2,005	2	2 Q1	092019	9	9 Jan
10	20190110	2019-01-10 00:00:00.000	10-01-2019	2,005	2,005	2	2 Q1	102019	10	10 Jan
11	20190111	2019-01-11 00:00:00.000	11-01-2019	2,005	2,005	2	2 Q1	112019	11	11 Jan
12	20190112	2019-01-12 00:00:00.000	12-01-2019	2,005	2,005	2	2 Q1	122019	12	12 Jan
13	20190113	2019-01-13 00:00:00.000	13-01-2019	2,005	2,005	2	2 Q1	012020	1	1 Feb
14	20190114	2019-01-14 00:00:00.000	14-01-2019	2,005	2,005	2	2 Q1	022020	2	2 Feb
15	20190115	2019-01-15 00:00:00.000	15-01-2019	2,005	2,005	2	2 Q1	032020	3	3 Feb
16	20190116	2019-01-16 00:00:00.000	16-01-2019	2,005	2,005	2	2 Q1	042020	4	4 Feb
17	20190117	2019-01-17 00:00:00.000	17-01-2019	2,005	2,005	2	2 Q1	052020	5	5 Feb
18	20190118	2019-01-18 00:00:00.000	18-01-2019	2,005	2,005	2	2 Q1	062020	6	6 Feb
19	20190119	2019-01-19 00:00:00.000	19-01-2019	2,005	2,005	2	2 Q1	072020	7	7 Feb
20	20190120	2019-01-20 00:00:00.000	20-01-2019	2,005	2,005	2	2 Q1	082020	8	8 Feb
21	20190121	2019-01-21 00:00:00.000	21-01-2019	2,005	2,005	2	2 Q1	092020	9	9 Feb
22	20190122	2019-01-22 00:00:00.000	22-01-2019	2,005	2,005	2	2 Q1	102020	10	10 Feb
23	20190123	2019-01-23 00:00:00.000	23-01-2019	2,005	2,005	2	2 Q1	112020	11	11 Feb
24	20190124	2019-01-24 00:00:00.000	24-01-2019	2,005	2,005	2	2 Q1	122020	12	12 Feb
25	20190125	2019-01-25 00:00:00.000	25-01-2019	2,005	2,005	2	2 Q1	012021	1	1 Mar
26	20190126	2019-01-26 00:00:00.000	26-01-2019	2,005	2,005	2	2 Q1	022021	2	2 Mar
27	20190127	2019-01-27 00:00:00.000	27-01-2019	2,005	2,005	2	2 Q1	032021	3	3 Mar
28	20190128	2019-01-28 00:00:00.000	28-01-2019	2,005	2,005	2	2 Q1	042021	4	4 Mar
29	20190129	2019-01-29 00:00:00.000	29-01-2019	2,005	2,005	2	2 Q1	052021	5	5 Mar
30	20190130	2019-01-30 00:00:00.000	30-01-2019	2,005	2,005	2	2 Q1	062021	6	6 Mar
31	20190131	2019-01-31 00:00:00.000	31-01-2019	2,005	2,005	2	2 Q1	072021	7	7 Mar
32	20190201	2019-02-01 00:00:00.000	01-02-2019	2,005	2,005	1	1 Q2	012021	1	1 Apr
33	20190202	2019-02-02 00:00:00.000	02-02-2019	2,005	2,005	1	1 Q2	022021	2	2 Apr
34	20190203	2019-02-03 00:00:00.000	03-02-2019	2,005	2,005	1	1 Q2	032021	3	3 Apr
35	20190204	2019-02-04 00:00:00.000	04-02-2019	2,005	2,005	1	1 Q2	042021	4	4 Apr
36	20190205	2019-02-05 00:00:00.000	05-02-2019	2,005	2,005	1	1 Q2	052021	5	5 Apr
37	20190206	2019-02-06 00:00:00.000	06-02-2019	2,005	2,005	1	1 Q2	062021	6	6 Apr
38	20190207	2019-02-07 00:00:00.000	07-02-2019	2,005	2,005	1	1 Q2	072021	7	7 Apr
39	20190208	2019-02-08 00:00:00.000	08-02-2019	2,005	2,005	1	1 Q2	082021	8	8 Apr
40	20190209	2019-02-09 00:00:00.000	09-02-2019	2,005	2,005	1	1 Q2	092021	9	9 Apr
41	20190210	2019-02-10 00:00:00.000	10-02-2019	2,005	2,005	1	1 Q2	102021	10	10 Apr
42	20190211	2019-02-11 00:00:00.000	11-02-2019	2,005	2,005	1	1 Q2	112021	11	11 Apr
43	20190212	2019-02-12 00:00:00.000	12-02-2019	2,005	2,005	1	1 Q2	122021	12	12 Apr
44	20190213	2019-02-13 00:00:00.000	13-02-2019	2,005	2,005	1	1 Q2	012022	1	1 May
45	20190214	2019-02-14 00:00:00.000	14-02-2019	2,005	2,005	1	1 Q2	022022	2	2 May
46	20190215	2019-02-15 00:00:00.000	15-02-2019	2,005	2,005	1	1 Q2	032022	3	3 May
47	20190216	2019-02-16 00:00:00.000	16-02-2019	2,005	2,005	1	1 Q2	042022	4	4 May
48	20190217	2019-02-17 00:00:00.000	17-02-2019	2,005	2,005	1	1 Q2	052022	5	5 May
49	20190218	2019-02-18 00:00:00.000	18-02-2019	2,005	2,005	1	1 Q2	062022	6	6 May
50	20190219	2019-02-19 00:00:00.000	19-02-2019	2,005	2,005	1	1 Q2	072022	7	7 May
51	20190220	2019-02-20 00:00:00.000	20-02-2019	2,005	2,005	1	1 Q2	082022	8	8 May
52	20190221	2019-02-21 00:00:00.000	21-02-2019	2,005	2,005	1	1 Q2	092022	9	9 May
53	20190222	2019-02-22 00:00:00.000	22-02-2019	2,005	2,005	1	1 Q2	102022	10	10 May
54	20190223	2019-02-23 00:00:00.000	23-02-2019	2,005	2,005	1	1 Q2	112022	11	11 May
55	20190224	2019-02-24 00:00:00.000	24-02-2019	2,005	2,005	1	1 Q2	122022	12	12 May
56	20190225	2019-02-25 00:00:00.000	25-02-2019	2,005	2,005	1	1 Q2	012023	1	1 Jun
57	20190226	2019-02-26 00:00:00.000	26-02-2019	2,005	2,005	1	1 Q2	022023	2	2 Jun
58	20190227	2019-02-27 00:00:00.000	27-02-2019	2,005	2,005	1	1 Q2	032023	3	3 Jun
59	20190228	2019-02-28 00:00:00.000	28-02-2019	2,005	2,005	1	1 Q2	042023	4	4 Jun
60	20190229	2019-02-29 00:00:00.000	29-02-2019	2,005	2,005	1	1 Q2	052023	5	5 Jun
61	20190301	2019-03-01 00:00:00.000	01-03-2019	2,005	2,005	4	4 Q3	012023	1	1 Sep
62	20190302	2019-03-02 00:00:00.000	02-03-2019	2,005	2,005	4	4 Q3	022023	2	2 Sep

Gambar 3.10 Data Tabel dim_date

6. Tabel Fakta Rental

Sebagai inti data warehouse, tabel fakta (fact_rental) menyimpan catatan transaksi penyewaan dan menghubungkan berbagai tabel dimensi melalui kolom foreign key. Data diambil dari tabel staging seperti rental, inventory, dan payment, yang kemudian melalui transformasi untuk melakukan lookup ke tabel dimensi serta penggabungan atribut kuantitatif seperti jumlah pembayaran. Tabel ini juga menambahkan kolom sales_id sebagai primary key menggunakan add sequence. Hasil akhirnya disimpan dalam tabel fact_rental. Gambar 3.11 menampilkan tampilan transformasi tf_fact_rental, sementara Gambar 3.12 menunjukkan hasil data dalam tabel fact_rental.



Gambar 3.11 Transformasi `tf_fact_rental`

	sk_payment_data	sk_rental_data	sk_return_data	sk_film	sk_customer	sk_staff	sk_store	amount	payment_data	rental_data	return_data	sales_id
1	20,070,430	20,050,524	20,050,520	334	480	1	2	4.99	2007-04-30 00:00:00.000	2005-05-24 00:00:00.000	2005-05-28 00:00:00.000	1
2	20,070,430	20,050,524	20,050,491	374	490	1	2	4.99	2007-04-30 00:00:00.000	2005-05-24 00:00:00.000	2005-06-01 00:00:00.000	2
3	20,070,430	20,050,524	20,050,603	537	334	2	1	2.99	2007-04-30 00:00:00.000	2005-05-24 00:00:00.000	2005-06-03 00:00:00.000	3
4	20,070,430	20,050,524	20,050,602	432	223	1	2	1.99	2007-04-30 00:00:00.000	2005-05-24 00:00:00.000	2005-06-02 00:00:00.000	4
5	20,070,430	20,050,524	20,050,527	615	549	1	1	0.99	2007-04-30 00:00:00.000	2005-05-24 00:00:00.000	2005-05-27 00:00:00.000	5
6	20,070,514	20,050,524	20,050,529	871	270	2	2	6	2007-05-14 00:00:00.000	2005-05-24 00:00:00.000	2005-05-29 00:00:00.000	6
7	20,070,430	20,050,524	20,050,527	511	240	2	1	5.99	2007-04-30 00:00:00.000	2005-05-24 00:00:00.000	2005-05-27 00:00:00.000	7
8	20,070,430	20,050,525	20,050,526	556	127	1	1	4.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-28 00:00:00.000	8
9	20,070,430	20,050,525	20,050,531	397	400	2	2	2.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-31 00:00:00.000	9
10	20,070,514	20,050,525	20,050,602	971	143	2	1	0.99	2007-05-14 00:00:00.000	2005-05-25 00:00:00.000	2005-06-02 00:00:00.000	10
11	20,070,430	20,050,525	20,050,530	546	262	2	1	4.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-30 00:00:00.000	11
12	20,070,514	20,050,525	20,050,530	499	335	1	1	1.99	2007-05-14 00:00:00.000	2005-05-25 00:00:00.000	2005-05-30 00:00:00.000	12
13	20,070,430	20,050,525	20,050,526	595	447	1	2	0.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-26 00:00:00.000	13
14	20,070,430	20,050,525	20,050,603	671	330	1	1	0.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-06-03 00:00:00.000	14
15	20,070,430	20,050,525	20,050,526	42	317	2	1	5.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-29 00:00:00.000	15
16	20,070,430	20,050,525	20,050,527	186	575	1	2	4.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-27 00:00:00.000	16
17	20,070,430	20,050,525	20,050,531	743	20	2	1	7.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-31 00:00:00.000	17
18	20,070,430	20,050,525	20,050,531	453	497	1	2	4.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-31 00:00:00.000	18
19	20,070,430	20,050,525	20,050,527	770	186	2	2	2.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-27 00:00:00.000	19
20	20,070,514	20,050,525	20,050,526	35	389	2	1	0.99	2007-05-14 00:00:00.000	2005-05-25 00:00:00.000	2005-05-26 00:00:00.000	20
21	20,070,430	20,050,525	20,050,526	142	310	1	2	4.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-26 00:00:00.000	21
22	20,070,514	20,050,525	20,050,526	871	419	1	1	0.99	2007-05-14 00:00:00.000	2005-05-25 00:00:00.000	2005-05-29 00:00:00.000	22
23	20,070,427	20,050,525	20,050,527	723	351	1	1	4.99	2007-04-27 00:00:00.000	2005-05-25 00:00:00.000	2005-05-27 00:00:00.000	23
24	20,070,430	20,050,525	20,050,527	864	28	2	2	5.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-27 00:00:00.000	24
25	20,070,430	20,050,525	20,050,531	993	372	1	1	0.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-31 00:00:00.000	25
26	20,070,430	20,050,525	20,050,530	272	302	2	2	4.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-30 00:00:00.000	26
27	20,070,430	20,050,525	20,050,526	189	213	2	1	2.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-26 00:00:00.000	27
28	20,070,514	20,050,525	20,050,530	136	45	2	2	4.99	2007-05-14 00:00:00.000	2005-05-25 00:00:00.000	2005-05-30 00:00:00.000	28
29	20,070,430	20,050,525	20,050,530	820	411	1	1	2.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-30 00:00:00.000	29
30	20,070,514	20,050,525	20,050,530	179	370	1	2	0.99	2007-05-14 00:00:00.000	2005-05-25 00:00:00.000	2005-05-30 00:00:00.000	30
31	20,070,430	20,050,525	20,050,525	619	211	1	2	1.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-25 00:00:00.000	31
32	20,070,430	20,050,525	20,050,527	368	273	1	1	10.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-27 00:00:00.000	32
33	20,070,514	20,050,525	20,050,529	124	597	2	1	4.99	2007-05-14 00:00:00.000	2005-05-25 00:00:00.000	2005-05-29 00:00:00.000	33
34	20,070,430	20,050,525	20,050,527	285	493	2	2	1.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-27 00:00:00.000	34
35	20,070,429	20,050,525	20,050,529	289	89	1	2	2.99	2007-04-29 00:00:00.000	2005-05-25 00:00:00.000	2005-05-29 00:00:00.000	35
36	20,070,430	20,050,525	20,050,529	95	535	1	2	4.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-29 00:00:00.000	36
37	20,070,430	20,050,525	20,050,491	440	303	1	1	1.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-06-01 00:00:00.000	37
38	20,070,430	20,050,525	20,050,531	976	208	2	1	2.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-31 00:00:00.000	38
39	20,070,428	20,050,525	20,050,527	580	414	1	1	0.99	2007-04-28 00:00:00.000	2005-05-25 00:00:00.000	2005-05-27 00:00:00.000	39
40	20,070,430	20,050,525	20,050,602	393	173	1	2	2.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-06-02 00:00:00.000	40
41	20,070,430	20,050,525	20,050,531	90	524	2	1	6.99	2007-04-30 00:00:00.000	2005-05-25 00:00:00.000	2005-05-31 00:00:00.000	41

Gambar 3.12 Data Tabel `fact_rental`

B. Implementasi CUBE dan OLAP

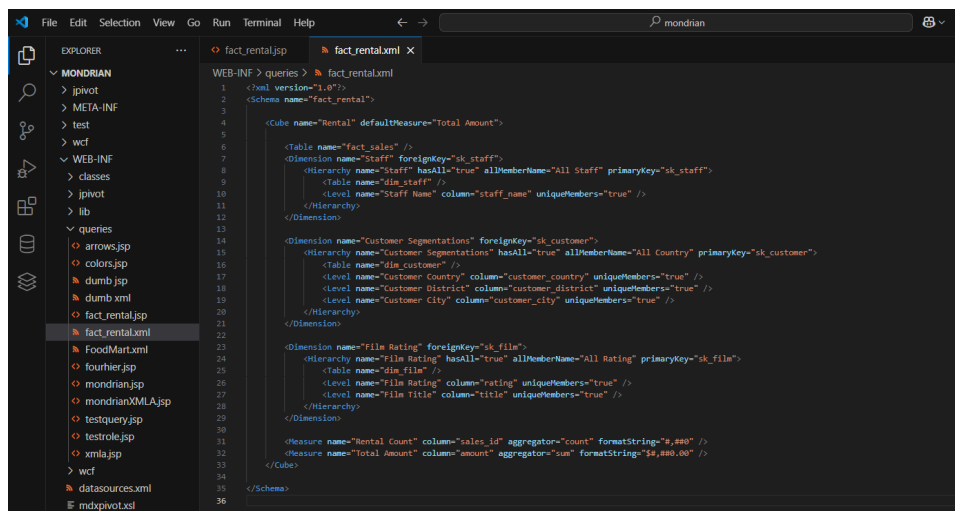
1. `fact_rental.xml`

File `fact_rental.xml` adalah file schema yang mendefinisikan struktur data multidimensi untuk analisis OLAP. File ini mendeskripsikan schema bernama `fact_rental`, yang mencakup cube, dimensi, hierarki, dan measures. Schema ini digunakan sebagai dasar oleh file JSP seperti `fact_rental.jsp` untuk mengeksekusi kueri OLAP.

Cube utama dalam file ini diberi nama `Rental`, dengan measure default `Total Amount`. Cube ini diambil dari tabel fakta bernama `fact_sales`, yang berisi data transaksi utama untuk dianalisis. File ini juga mendefinisikan beberapa dimensi yang relevan untuk analisis data. Dimensi pertama adalah `Staff`, yang dihubungkan dengan tabel `dim_staff` melalui kunci asing `sk_staff`. Dimensi ini memiliki hierarki yang memungkinkan analisis berdasarkan nama staf (`Staff Name`). Dimensi kedua adalah `Customer Segmentations`, yang diambil dari tabel `dim_customer` dengan kunci asing `sk_customer`. Dimensi ini memiliki level hierarki seperti `Customer Country` dan `Customer City`, yang memungkinkan analisis berdasarkan lokasi geografis pelanggan. Dimensi ketiga adalah `Film Rating`, yang dihubungkan dengan tabel `dim_film` menggunakan kunci asing `sk_film`. Dimensi ini mencakup level seperti `Film Rating` dan `Film Title`, sehingga analisis dapat dilakukan berdasarkan kategori atau judul film.

Selain dimensi, file ini juga mendefinisikan dua measures utama. Measure pertama adalah `Rental Count`, yang menghitung jumlah transaksi berdasarkan kolom `sales_id` dengan metode agregasi `count`. Measure kedua adalah `Total Amount`, yang menjumlahkan nilai pendapatan dari kolom `amount` menggunakan metode agregasi `sum`. Measure ini dilengkapi dengan format angka untuk menampilkan hasil dalam format mata uang.

Struktur dimensi dan measures dihubungkan melalui relasi antar tabel menggunakan kunci asing, seperti sk_staff, sk_customer, dan sk_film. Hubungan ini memungkinkan data multidimensi dianalisis secara menyeluruh. File schema ini memainkan peran penting dalam menyediakan kerangka kerja untuk analisis data OLAP yang dilakukan di file JSP.



Gambar 3.13 fact_rental.xml

2. fact_rental.jsp

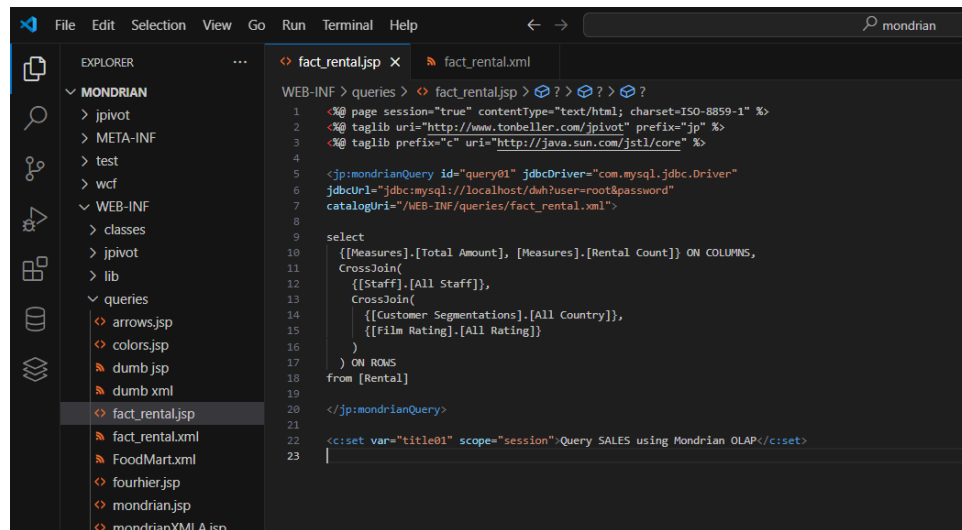
File fact_rental.jsp merupakan file JSP yang digunakan untuk menampilkan data multidimensi dengan memanfaatkan framework Mondrian OLAP. Di bagian awal file, terdapat konfigurasi halaman dengan encoding ISO-8859-1 yang memastikan karakter ditampilkan dengan benar. Selain itu, file ini juga menggunakan pustaka tag (taglib) JPivot untuk mengintegrasikan fitur analisis data OLAP ke dalam aplikasi berbasis web.

File ini memuat koneksi ke database melalui tag `<jp:mondrianQuery>`. Koneksi dilakukan dengan menggunakan driver MySQL (`com.mysql.jdbc.Driver`) dan URL `jdbc:mysql://localhost/dwh?user=root&password=`, yang menunjukkan bahwa file ini terhubung ke database bernama dwh di server lokal dengan user root. Selain itu, file ini merujuk pada file schema fact_rental.xml sebagai sumber definisi struktur data multidimensi.

Pada bagian inti, file ini menjalankan kueri OLAP menggunakan bahasa MDX (Multi-Dimensional Expressions). Kueri ini mendefinisikan dua measures, yaitu [Total Amount] untuk total pendapatan dan [Rental Count] untuk jumlah transaksi. Kedua measures ini ditampilkan pada kolom, sedangkan baris diisi dengan kombinasi dimensi Staff, Customer Segmentations, dan Film Rating. Kombinasi ini menggunakan fungsi CrossJoin untuk menghasilkan analisis multidimensi yang lebih dalam. Data yang diambil berasal dari cube Rental yang sudah didefinisikan dalam file fact_rental.xml.

Di bagian akhir, file ini menyimpan variabel sesi bernama title01 menggunakan tag `<c:set>` dengan nilai string "Query SALES using Mondrian OLAP". Variabel ini bisa

digunakan untuk memberikan konteks tambahan atau keterangan pada antarmuka pengguna. Secara keseluruhan, file ini bertindak sebagai antarmuka pengguna untuk mengeksekusi dan menampilkan hasil analisis data berbasis OLAP.



Gambar 3.14 fact_rental.jsp

3. Output Mondrian

a.) Tampilan Awal

Query SALES using Mondrian OLAP



			Measures	
Staff	Customer Segmentations	Film Rating	Total Amount	Rental Count
All Staff	All Country	All Rating	\$59,398.30	16,044

Slicer:

[back to index](#)

Gambar 3.15 Tampilan Awal Mondrian

b.) Dimensi Staff

Query SALES using Mondrian OLAP



			Measures	
Staff	Customer Segmentations	Film Rating	Total Amount	Rental Count
All Staff	All Country	All Rating	\$59,398.30	16,044
Jon Stephens	All Country	All Rating	\$29,678.18	8,004
Mike Hillyer	All Country	All Rating	\$29,720.12	8,040

Slicer:

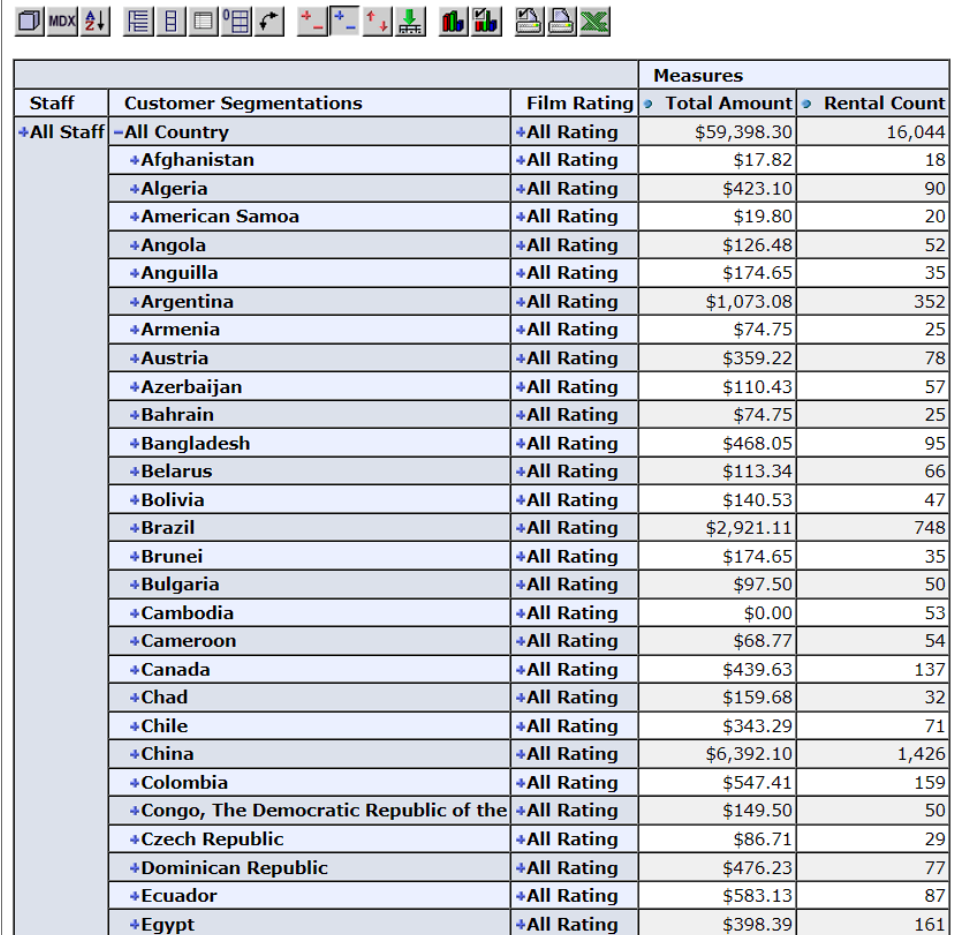
[back to index](#)

Gambar 3.16 Tampilan Dimensi Store Mondrian

c.) Dimensi Customer Demographics

1.) Customer Country

Query SALES using Mondrian OLAP



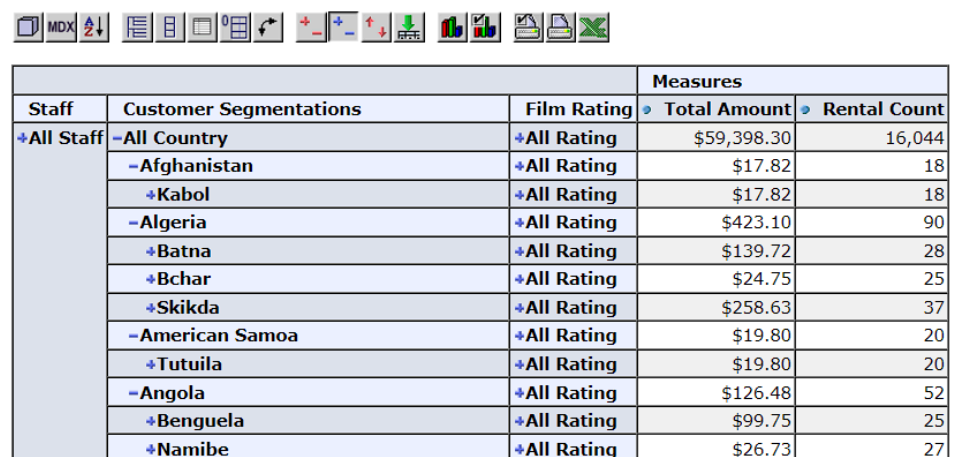
The screenshot shows the Mondrian OLAP interface with a toolbar at the top. The main table displays the results of a query for 'Customer Country'. The table has five columns: 'Staff', 'Customer Segmentations', 'Film Rating', 'Total Amount', and 'Rental Count'. The data is filtered by 'All Staff' and 'All Country'. The table lists 30 countries with their respective 'All Rating' and financial data.

Staff	Customer Segmentations	Film Rating	Total Amount	Rental Count
✦All Staff	✦All Country	✦All Rating	\$59,398.30	16,044
	✦Afghanistan	✦All Rating	\$17.82	18
	✦Algeria	✦All Rating	\$423.10	90
	✦American Samoa	✦All Rating	\$19.80	20
	✦Angola	✦All Rating	\$126.48	52
	✦Anguilla	✦All Rating	\$174.65	35
	✦Argentina	✦All Rating	\$1,073.08	352
	✦Armenia	✦All Rating	\$74.75	25
	✦Austria	✦All Rating	\$359.22	78
	✦Azerbaijan	✦All Rating	\$110.43	57
	✦Bahrain	✦All Rating	\$74.75	25
	✦Bangladesh	✦All Rating	\$468.05	95
	✦Belarus	✦All Rating	\$113.34	66
	✦Bolivia	✦All Rating	\$140.53	47
	✦Brazil	✦All Rating	\$2,921.11	748
	✦Brunei	✦All Rating	\$174.65	35
	✦Bulgaria	✦All Rating	\$97.50	50
	✦Cambodia	✦All Rating	\$0.00	53
	✦Cameroon	✦All Rating	\$68.77	54
	✦Canada	✦All Rating	\$439.63	137
	✦Chad	✦All Rating	\$159.68	32
	✦Chile	✦All Rating	\$343.29	71
	✦China	✦All Rating	\$6,392.10	1,426
	✦Colombia	✦All Rating	\$547.41	159
	✦Congo, The Democratic Republic of the	✦All Rating	\$149.50	50
	✦Czech Republic	✦All Rating	\$86.71	29
	✦Dominican Republic	✦All Rating	\$476.23	77
	✦Ecuador	✦All Rating	\$583.13	87
	✦Egypt	✦All Rating	\$398.39	161

Gambar 3.17 Tampilan Dimensi Customer Country Mondrian

2.) Customer District

Query SALES using Mondrian OLAP



The screenshot shows the Mondrian OLAP interface with a toolbar at the top. The main table displays the results of a query for 'Customer District'. The table has five columns: 'Staff', 'Customer Segmentations', 'Film Rating', 'Total Amount', and 'Rental Count'. The data is filtered by 'All Staff' and 'All Country'. The table lists 13 districts with their respective 'All Rating' and financial data.

Staff	Customer Segmentations	Film Rating	Total Amount	Rental Count
✦All Staff	✦All Country	✦All Rating	\$59,398.30	16,044
	✦Afghanistan	✦All Rating	\$17.82	18
	✦Kabul	✦All Rating	\$17.82	18
	✦Algeria	✦All Rating	\$423.10	90
	✦Batna	✦All Rating	\$139.72	28
	✦Bchar	✦All Rating	\$24.75	25
	✦Skikda	✦All Rating	\$258.63	37
	✦American Samoa	✦All Rating	\$19.80	20
	✦Tutuila	✦All Rating	\$19.80	20
	✦Angola	✦All Rating	\$126.48	52
	✦Benguela	✦All Rating	\$99.75	25
	✦Namibe	✦All Rating	\$26.73	27

Gambar 3.18 Tampilan Dimensi Customer District Mondrian

3.) Customer City

Query SALES using Mondrian OLAP


Staff	Customer Segmentations	Film Rating	Total Amount	Rental Count
All Staff	All Country	All Rating	\$59,398.30	16,044
	Afghanistan	All Rating	\$17.82	18
	Kabol	All Rating	\$17.82	18
	Kabul	All Rating	\$17.82	18
	Algeria	All Rating	\$423.10	90
	Batna	All Rating	\$139.72	28
	Batna	All Rating	\$139.72	28
	Bchar	All Rating	\$24.75	25
	Bchar	All Rating	\$24.75	25
	Skikda	All Rating	\$258.63	37
	Skikda	All Rating	\$258.63	37
	American Samoa	All Rating	\$19.80	20
	Tutuila	All Rating	\$19.80	20
	Tafuna	All Rating	\$19.80	20
	Angola	All Rating	\$126.48	52
	Benguela	All Rating	\$99.75	25
	Benguela	All Rating	\$99.75	25
	Namibe	All Rating	\$26.73	27
	Namibe	All Rating	\$26.73	27
	Anguilla	All Rating	\$174.65	35
	Argentina	All Rating	\$1,073.08	352
	Buenos Aires	All Rating	\$715.84	276
	Crdoba	All Rating	\$24.75	25
	Santa F	All Rating	\$188.73	27
	Tucumn	All Rating	\$143.76	24

Gambar 3.19 Tampilan Dimensi Customer City Mondrian

d.) Dimensi Film

1.) Film Rating

Query SALES using Mondrian OLAP



Staff	Customer Segmentations	Film Rating	Measures	
			Total Amount	Rental Count
+All Staff	+All Country	-All Rating	\$59,398.30	16,044
		+G	\$10,356.34	2,773
		+NC-17	\$12,096.47	3,293
		+PG	\$11,780.39	3,212
		+PG-13	\$13,303.70	3,585
		+R	\$11,861.40	3,181


Slicer:

[back to index](#)

Gambar 3.20 Tampilan Dimensi Film Category Mondrian

2.) Film Tittle

Query SALES using Mondrian OLAP



		Measures	
Staff	Customer Segmentations	Film Rating	
+All Staff	+All Country	-All Rating	\$59,398.30 16,044
		-G	\$10,356.34 2,773
		Ace Goldfinger	\$33.93 7
		Affair Prejudice	\$79.79 23
		African Egg	\$46.88 12
		Alamo Videotape	\$75.76 24
		Amistad Midsummer	\$59.80 21
		Angels Life	\$61.81 22
		Annie Identity	\$77.79 22
		Armageddon Lost	\$74.79 21
		Atlantis Cause	\$68.77 24
		Autumn Crow	\$47.90 11
		Baked Cleopatra	\$60.85 16
		Ballroom Mockingbird	\$46.92 8
		Barbarella Streetcar	\$112.75 25
		Barefoot Manchurian	\$59.82 18
		Beach Heartbreakers	\$64.83 17
		Beauty Grease	\$85.79 22
		Bird Independence	\$42.89 11
		Birds Perdition	\$69.84 18
		Blanket Beverly	\$43.88 12
		Blood Argonauts	\$43.90 10
		Blues Instinct	\$69.80 21
		Bonnie Holocaust	\$80.84 16
		Borrowers Bedazzled	\$86.79 22
		Breakfast Goldfinger	\$60.85 17
		Bride Intrigue	\$59.83 19
		Bugsy Song	\$31.89 11

Gambar 3.21 Tampilan Dimensi Film Tittle Mondrian

BAB IV : PENUTUPAN

A. Kesimpulan

Proyek pengembangan Data Warehouse dan OLAP untuk studi kasus DVD Rental berhasil diselesaikan dengan hasil yang memuaskan. Seluruh tahapan proyek telah dilakukan secara sistematis, memberikan pengalaman berharga dalam pengelolaan data dan mendukung pengambilan keputusan berbasis data.

Tahapan awal dimulai dengan perancangan skema Data Warehouse menggunakan pendekatan Star Schema. Desain ini memfasilitasi hubungan antara tabel fakta dan tabel dimensi, memastikan struktur data yang sederhana dan mudah diakses untuk analisis lebih lanjut. Skema ini dirancang untuk mengakomodasi berbagai aspek bisnis DVD Rental, seperti transaksi, pelanggan, dan kategori film.

Proses selanjutnya adalah ETL (Extract, Transform, Load), yang dilakukan menggunakan Pentaho Data Integration. Tahap ini mencakup ekstraksi data dari sumber asli, transformasi data agar sesuai dengan skema Data Warehouse, serta pemuatan data ke dalam sistem. Dengan memanfaatkan alat ini, proses ETL berjalan efisien dan memastikan kualitas data yang tinggi, sehingga mendukung analisis yang akurat.

Tahap akhir adalah implementasi OLAP Cube menggunakan Mondrian OLAP. Cube yang dihasilkan memungkinkan analisis data multidimensi yang mendalam, mencakup performa toko, wilayah pelanggan, hingga kategori film yang paling diminati. Hasil akhir ini memberikan wawasan strategis yang dapat digunakan untuk mendukung pengambilan keputusan bisnis secara lebih terinformasi.

Secara keseluruhan, proyek ini menunjukkan pentingnya integrasi teknologi dalam pengelolaan data. Kombinasi Data Warehouse, ETL, dan OLAP Cube menghasilkan sistem yang tidak hanya efisien dalam menyimpan data tetapi juga powerful dalam memberikan insight strategis.

B. Saran

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

- 1.) Mengujicoba dengan dataset yang lebih besar untuk mengukur seberapa baik kinerja Data Warehouse dan OLAP dalam skala yang lebih besar.
- 2.) Menggunakan alat visualisasi lain seperti Tableau atau Power BI untuk memperbaiki tampilan dan kualitas analisis.
- 3.) Mengembangkan analisis dengan menambah dimensi atau ukuran baru, misalnya analisis yang lebih mendalam berdasarkan waktu transaksi.

- 4.) Menambahkan otomatisasi untuk pembaruan data di Data Warehouse, agar analisis bisa dilakukan secara langsung dan real-time.

LAMPIRAN

Lampiran 1: Link GitHub

https://github.com/rafieahzaghaisan/Konversi_DWO_22082010192_Rafie-Ahza-Ghaisan