

PERANCANGAN DATA WAREHOUSE UNTUK PREDIKSI PENJUALAN PRODUK PADA ORBA EXPRESS MENGGUNAKAN PENTAHO

I Putu Agus Eka Pratama¹⁾, Ni Putu Nirmala Dewi Widhiasih²⁾

¹Fakultas Teknik, Universitas Udayana
email: eka.pratama@unud.ac.id

²Fakultas Teknik, Universitas Udayana
email: lalanirmala1911@gmail.com

Abstract

The *Data Warehouse* is a large data *Storage* facility or something called a *Data Warehouse*. The human work that requires so much data makes it harder to do it by hand. Such large amounts of data then hinder one from managing the data and predicting the products that are so popular by consumers that they decide to store them in the *Data Warehouse*. The *Data Warehouse* is the solution to the problem to predict a sale of a product. The *Data Warehouse* is a system that can extract, clean, adapt, and send a source of data into a data *Storage* or *Data Warehouse* which then implement through tools *Pentaho* in order to be able to analyze data. The purpose of the *Data Warehouse* design was to be able to mitigate shop owners' efforts on decision-making, like deciding what products sell to stock stores so they can reduce the stores' built-up.

Keywords: *Data Warehouse, Storage, Pentaho.*

1. PENDAHULUAN

1.1 Latar Belakang

Penduduk di Indonesia setiap tahunnya semakin meningkat begitu pula dengan kebutuhan yang diperlukan salah satunya membangun rumah impian. Setiap orang dalam membangun rumah impian yang nyaman membutuhkan furniture pendukung di dalamnya. Industry mebel merupakan salah satu industry yang dapat memenuhi kebutuhan masyarakat. Kebutuhan furniture rumah tangga dengan desain beragam yang memiliki nilai artistic dapat memberikan kenyamanan kepada setiap orang dalam melakukan aktifitas. Orba Express Furniture adalah salah satu industry mebel yang berlokasi di jalan Gunung Salak No.115, Padang Sambian Kelod, Denpasar, Bali, Indonesia. Orba Express merupakan bisnis yang melakukan ekspor dan impor karena banyaknya minat furniture mebel baik di dalam negeri maupun luar negeri. Orba Express Furnitur memproduksi berbagai olahan produk yang berbahan baku berupa kayu jati, rotan, mahoni, dll. Bahan baku tersebut berasal dari pulau jawa. Permasalahan yang sedang terjadi oleh Orba Express Furniture yaitu sulitnya memprediksi produk yang paling diminati konsumen dimana ketika konsumen ingin membeli produk di toko tetapi produk tersebut tidak dijual lagi dan banyaknya stok produk

yang tersisa disebabkan oleh kurangnya pengelolaan data sehingga membuat data tidak terstruktur. Tujuan perancangan data warehouse untuk Orba Express Furniture agar dapat menyimpan data yang berasal dari berbagai sumber ke dalam sebuah tempat penyimpanan sehingga pengguna dapat dengan mudah melakukan pencarian data dan memudahkan dalam pengambilan keputusan serta analisis.

1.2 Rencana Penelitian

Rencana pemecahan masalah yang digunakan untuk mengatasi permasalahan pada Orba Express Furniture adalah dengan penerapan teknologi data warehouse dengan menggunakan MySQL dan tools Pentaho.

1.3 Tujuan Penelitian

Tujuan dari pembuatan paper dengan judul "Perancangan Data Warehouse untuk Prediksi Penjualan Produk Pada Orba Express Menggunakan Pentaho" adalah untuk membuat perancangan data warehouse agar dapat memprediksi penjualan produk pada Orba Express Furniture sehingga dapat digunakan untuk pengambilan keputusan.

1.4 Tinjauan Pustaka

1.4.1 Data Warehouse

Data Warehouse adalah gudang data yang menyimpan beragam data baik data terstruktur maupun data tidak terstruktur. Data Warehouse dapat mengirim sumber data ke dalam penyimpanan data dengan tujuan pengambilan keputusan sehingga memudahkan pekerjaan manusia. Sumber data yang di dapat bisa berasal dari database, catatan, dan data berupa word, excel, dll.

Data Warehouse memiliki karakteristik yang membedakannya dari database lainnya. Buku yang berjudul Data Warehousing Fundamentals terdapat empat karakteristik dari Data Warehouse yang dijelaskan sebagai berikut:

a. Subject Oriented

Data Warehouse mengorganisasikan subjek utama perusahaan/organisasi bukan pada area aplikasi utama. Hal ini menyebabkan data warehouse hanya mengorientasikan pada kebutuhan untuk menyimpan data pendukung keputusan dari pada aplikasi yang berorientasi data. Data warehouse hanya mengandung data yang dibutuhkan untuk fungsi yang berhubungan dengan sebagian aplikasi.

b. Data yang Terintegrasi

Data di dalam data warehouse akan memiliki perbedaan database, file dan segmentasi, dikarenakan sumber data untuk data warehouse diambil dari aplikasi – aplikasi yang berbeda sehingga flatform dan sistem operasi yang digunakan akan memunculkan perbedaan pada tampilan file, tabel, representasi kode karakter, dan juga penamaan file. Oleh karena itu, data yang akan digunakan untuk data warehouse harus melewati beberapa proses yaitu transformasi, konsolidasi, dan integrasi dengan sumber data lainnya.

c. Time Variant

Sistem operasional merupakan tempat data disimpan yang mengandung nilai saat ini saja, namun beberapa data yang sifatnya masa lalu, data warehouse mengandung data historis, tidak hanya nilai saat ini. Data disimpan sebagai gambaran masa lalu dan periode saat ini. Setiap struktur data dalam data warehouse adalah mengizinkan untuk menganalisis masa lalu, menghubungkan informasi saat ini, dan memprediksi masa depan.

d. Non Volatile

Data di dalam data warehouse tidak dirancang untuk menjalankan bisnis setiap harinya sehingga dalam data warehouse tidak ada proses pembaharuan data untuk setiap kali proses transaksi berjalan.(Dharmayanti, Bachtiar, and Heryandi 2013)

1.4.2 Data Multi Dimensi

Data multi dimensi merupakan dasar dari pembuatan data warehouse yang dilihat dari berbagai sudut pandang yang berbeda. Kebutuhan user yang menjadi faktor penentu untuk merancang model dimensi data warehouse sehingga perlu dirancang dengan sedemikian rupa agar mudah dipertahankan dan beradaptasi dari segala perubahan yang terjadi. Model data multi dimensi terdiri dari tabel dimensi dan tabel fakta.

Tabel dimensi adalah entitas bisnis dari suatu enterprise. Tabel dimensi umumnya berisi data keterangan, dimana data tersebut jarang sekali mengalami perubahan.

Tabel fakta merupakan sebuah tabel yang menjelaskan tentang transaksi bisnis dari suatu enterprise biasanya disebut tabel detail. Tabel fakta umumnya berisi data yang berkaitan langsung dengan proses bisnisnya.

Skema dimensional model adalah beberapa skema yang biasa digunakan untuk merancang suatu data warehouse dalam proses OLAP.

1.4.3 MySQL

MySQL adalah sebuah perangkat lunak atau software sistem manajemen basis data (DBMS) yang multi user. MySQL salah satu turunan dalam database sejak lama, yaitu SQL (Structured Query Language). SQL adalah sebuah konsep pengoperasian database untuk pemilihan, seleksi dan pemasukan data yang dapat melakukan pengoperasian data dengan mudah dan cepat secara otomatis, software ini merupakan aplikasi yang bersifat open source sehingga bebas menginstall, menggunakan, serta mengatur dan mengubah software tersebut sesuai dengan kebutuhan.

Menurut kustiyahningsih(2011:145), “MySQL adalah sebuah basis data yang mengandung satu atau jumlah tabel. Tabel terdiri atas sejumlah baris dan setiap baris mengandung satu atau sejumlah tabel”

sedangkan tipe data MySQL adalah data yang terdapat dalam sebuah tabel berupa field – field yang berisi nilai dari data tersebut. Nilai data dalam fiels memiliki tipe sendiri – sendiri.(Firman et al. 2016)

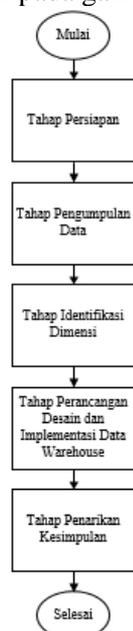
1.4.4 Pentaho

Pentaho Data Integration (PDI) atau kettle adalah software dari pentaho yang dapat digunakan untuk proses ETL (Extraction, Transformation, dan Loading). PDI dapat digunakan untuk migrasi data, membersihkan data, loading dari file ke database atau sebaliknya dalam volume besar. PDI menyediakan graphical user interface dan drag-drop komponen yang memudahkan user. Elemen utama dari PDI adalah Transformation dan job. Transformation adalah sekumpulan intruksi untuk merubah input menjadi output yang diinginkan (input-proses-output). Sedangkan job adalah kumpulan intruksi untuk menjalankan transformation. Ada tiga komponen dalam PDI yaitu spoon, pan, dan kitchen. Spoon adalah user interface untuk membuat job dan transformation. Pan adalah tools yang berfungsi membaca, merubah, dan menulis data. Sedangkan kitchen adalah program yang mengeksekusi job.(Wibisono 2014)

2. METODE PENELITIAN

2.1 Rancangan Metode

Proses pembuatan data warehouse pada Orba Express menggunakan beberapa tahapan yang digambarkan pada gambar 2.1.



Gambar 2.1 Rancangan Metode

Gambar 2.1 merupakan gambar dari rancangan metode penelitian dalam pembuatan data warehouse. Gambar tersebut menjelaskan beberapa tahapan yang dijelaskan sebagai berikut.

- 1) Tahap Persiapan
Tahap ini merupakan tahap awal penelitian yaitu dengan menentukan latar belakang, rencana penelitian, dan tujuan penelitian.
- 2) Tahap Pengumpulan Data
Tahap ini merupakan proses pengumpulan data dengan beberapa cara seperti observasi, wawancara, dan studi literature.
- 3) Tahap Identifikasi Dimensi
Tahap ini merupakan tahap menentukan data multi dimensi yang dilihat dari berbagai sudut pandang yang berbeda sesuai kebutuhan user.
- 4) Tahap Perancangan Desain dan Implementasi Data Warehouse
Tahap ini merupakan tahap mendesain arsitektur fisik dan logika data warehouse. Data yang didapat dari berbagai sumber akan diproses dengan konsep ETL (Extracts, Transformation, Loading).
- 5) Tahap Penarikan Kesimpulan
Tahap ini merupakan tahap akhir yaitu implementasi data warehouse sesuai dengan kebutuhan user. Dari hasil perancangan desain dan implementasi yang telah dilakukan, maka hasil akhir dari penelitian ini merupakan penarikan kesimpulan.(Pendukung et al. 2013)

3. HASIL DAN PEMBAHASAN

3.1 Persiapan Sumber Data

Persiapan sumber data dari Orba Express furniture belum memiliki sistem yang menggunakan database dalam menangani transaksi operasionalnya. Permasalahan tersebut menyebabkan kesulitan dalam menganalisis data dan menghambat pengambilan keputusan karena semua transaksi tidak dicatat secara lengkap. Oleh karena itu, dibuatkan sebuah database yang menangani operasional toko sehingga mempermudah implementasi data warehouse selanjutnya. Sumber data yang telah ada masih berbentuk file excel yang nantinya file tersebut menggunakan DBMS MySQL.

3.2 Rancangan Data Warehouse

3.2.1 Analisis Kebutuhan

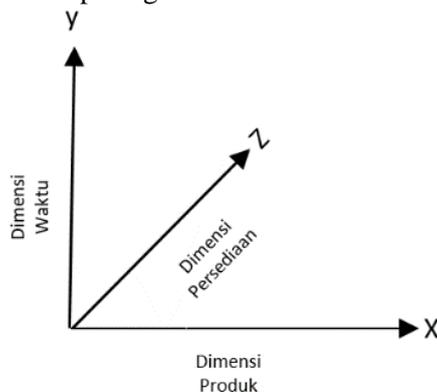
Analisis kebutuhan informasi pada perancangan data warehouse prediksi penjualan ini menggunakan data dari orba express furniture yang berisi informasi yaitu Nama Produk yang terjual, jumlah produk yang terjual, harga satuan produk, total harga yang terjual, dan tanggal produk yang terjual.

3.2.2 Memilih Grain

Grain merupakan proses data yang dapat dilakukan analisis setelah mengetahui proses bisnis yang ada. Hasil proses ini akan digambarkan sebagai tabel yang berisikan fakta. Grain yang dihasilkan meliputi waktu, nama produk terjual, jenis produk, kategori produk, dan jumlah produk terjual. (Khotimah and - 2016)

3.2.3 Identifikasi Dimensi

Identifikasi dimensi yaitu dimensi yang berhubungan dengan tabel fakta. Hasil identifikasi dimensi dipresentasikan ke dalam bentuk koordinat yang ditunjukkan ke dalam tiga buah sumbu: x, y, dan z yang digambarkan pada gambar 2.2.



Gambar 3.1 Multi Dimensi

Gambar 3.1 merupakan gambar multi dimensi dari dasar pembuatan data warehouse. Gambar tersebut menjelaskan beberapa dimensi berdasarkan kebutuhan user yang dijelaskan pada tabel fakta.

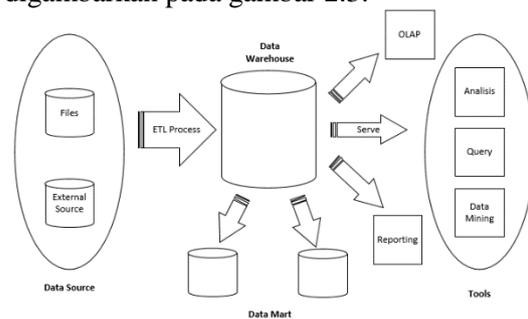
3.2.4 Memilih Fakta

Dimensi yang digunakan untuk melihat data penjualan adalah dimensi waktu, dimensi produk, dan dimensi persediaan. Masing – masing data dimensi tersebut membentuk tabel yang nantinya akan memiliki field sebagai berikut.

- Tabel Waktu berisi field: id_waktu, tanggal-bulan-tahun.
- Tabel Produk berisi field: id_produk, nama_produk, jenis_produk, dan harga_satuan_produk.
- Tabel Persediaan berisi field: id_produk, jumlah_produk.

3.2.5 Arsitektur Data Warehouse

Arsitektur data warehouse merupakan tahapan proses pengumpulan beberapa sumber yang terpisah kemudian disatukan kedalam satu tempat penyimpanan yang bertujuan untuk memudahkan proses load data ke dalam data mart. Arsitektur data warehouse digambarkan pada gambar 2.3.



Gambar 3.2 Arsitektur Data Warehouse

Gambar 3.2 merupakan arsitektur data warehouse prediksi penjualan produk pada Orba Express Furniture. Arsitektur data warehouse meliputi alat untuk mengekstrak data dari berbagai sumber data baik eksternal maupun dalam bentuk file yang kemudian mengalami proses ETL (Extraction, Transformation, Loading).

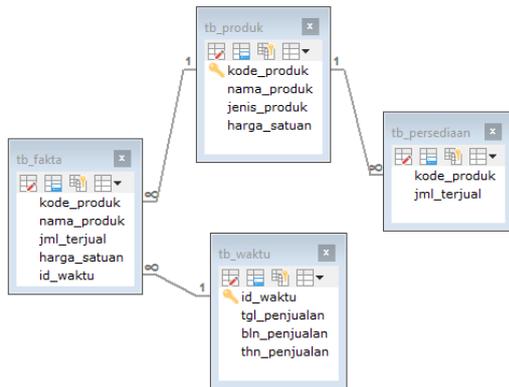
Jika telah masuk ke dalam data warehouse lalu dibagi menjadi sub bagian yang disebut data mart yang menghasilkan output berupa data produk yang tidak laku terjual dan data produk yang laku terjual.

Data dari data warehouse kemudian akan melalui proses OLAP (Online Analytical Processing) yaitu proses mempercepat pengambilan data dengan menggunakan data multi dimensi yang dipresentasikan sebagai x,y,z yaitu terdiri dari dimensi waktu, dimensi produk, dan dimensi persediaan.

3.2.6 Skema Data Warehouse

Skema data warehouse adalah langkah setelah melakukan analisis dan identifikasi tabel dimensi dengan

menggunakan tabel fakta. Skema database ini dibuat dengan menggunakan DBMS MySQL. Skema database data warehouse digambarkan pada gambar 3.3.

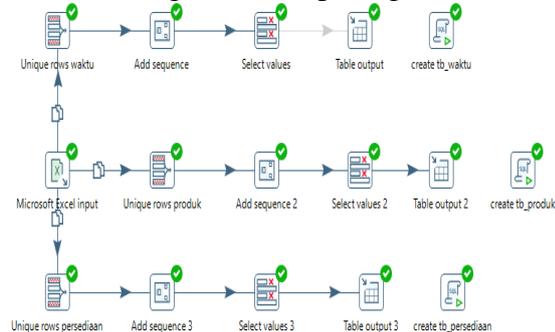


Gambar 3.3 Skema Database

Gambar 3.3 merupakan gambar skema database. Terdapat tb_fakta, tb_produk, tb_persediaan, dan tb_waktu yang saling berhubungan.

3.2.7 Proses ETL

Proses ETL adalah proses setelah membuat database pada DBMS MySQL lalu kemudian dari database operasional yang telah dibuat ke database data warehouse dengan menggunakan tools pentaho. Skema proses ETL yang dilakukan digambarkan pada gambar 3.4.



Gambar 3.4 Proses ETL

Gambar 3.4 merupakan gambar proses ETL, dimana data bersumber dari file excel yang membentuk tabel waktu, tabel produk, dan tabel persediaan yang akan langsung terbentuk pada DBMS MySQL. Penjelasan proses ETL adalah sebagai berikut.

1) Unique Rows

Unique rows merupakan tahap awal untuk memasukan field sesuai tabel fakta yang

telah dibuat, seperti tb_waktu berisi field tanggal, bulan, tahun, tb_produk berisi field kode produk, nama produk, jenis produk, harga satuan, dan tb_persediaan berisi field kode produk, jumlah terjual.

2) Add Sequence

Add sequence merupakan tahap memasukan id yang akan digunakan sebagai primary key, seperti kode_produk dan id_waktu. tahap ini juga memasukan koneksi database.

3) Select Value

Select value merupakan tahap memberikan penamaan untuk tabel yang akan terbuat pada database seperti kode produk :
kode_produk, nama produk :
nama_produk, jenis produk :
jenis_produk, harga satuan :
harga_satuan.

4) Table Output

Table output merupakan tahap akhir yang keluarannya berupa tabel yang dibentuk pada DBMS MySQL.

Hasil skema dari proses ETL terdiri dari tb_waktu, tb_produk, dan tb_persediaan yang digambarkan sebagai berikut.

```
MariaDB [db_penjualan]> SELECT *FROM tb_waktu;
```

| id_waktu | tgl_penjualan | bln_penjualan | thn_penjualan |
|----------|---------------|---------------|---------------|
| 1 | | 1 Januari | 2018 |
| 2 | | 8 Januari | 2018 |
| 3 | | 9 Januari | 2018 |
| 4 | | 17 Januari | 2018 |
| 5 | | 21 Januari | 2018 |
| 6 | | 30 Januari | 2018 |
| 7 | | 1 Februari | 2018 |
| 8 | | 10 Februari | 2018 |
| 9 | | 18 Februari | 2018 |
| 10 | | 22 Februari | 2018 |
| 11 | | 7 Maret | 2018 |
| 12 | | 8 Maret | 2018 |
| 13 | | 13 Maret | 2018 |
| 14 | | 20 Maret | 2018 |
| 15 | | 26 Maret | 2018 |
| 16 | | 10 April | 2018 |
| 17 | | 19 April | 2018 |
| 18 | | 20 April | 2018 |
| 19 | | 26 April | 2018 |
| 21 | | 3 Mei | 2018 |
| 22 | | 7 Mei | 2018 |
| 23 | | 8 Mei | 2018 |
| 24 | | 18 Mei | 2018 |
| 25 | | 30 Mei | 2018 |
| 27 | | 2 Juni | 2018 |

Gambar 3.5 Tampilan tb_waktu

Gambar 3.5 merupakan tampilan tb_waktu yang berisi id_waktu, tgl_penjualan, dan tahun_penjualan dari bulan januari sampai bulan desember tahun 2018.

```
MariaDB [db_penjualan]> SELECT *FROM tb_produk;
```

| kode_produk | nama_produk | jenis_produk | harga_satuan |
|-------------|--------------------------|-------------------------|--------------|
| PRD - 1183 | Bed side 1 laci | pendamping tempat tidur | 500000 |
| PRD - 1184 | Bed 160 x 200 | Tempat tidur | 2000000 |
| PRD - 1185 | Ayunan | Ayunan Kayu | 2500000 |
| PRD - 1186 | Set Kitchen | Dapur | 8000000 |
| PRD - 1187 | Meja Belajar | Meja | 1500000 |
| PRD - 1188 | Meja Dapur | Meja | 1000000 |
| PRD - 1189 | Meja Kaca | Meja | 1500000 |
| PRD - 1190 | Kursi Rotan | Kursi | 500000 |
| PRD - 1191 | Kursi Bar | Kursi | 600000 |
| PRD - 1192 | Kursi Cafe | Kursi | 600000 |
| PRD - 1193 | Tempat Payung | Tempat Payung | 1500000 |
| PRD - 1194 | Bed 180 x 200 | Tempat Tidur | 3000000 |
| PRD - 1195 | Washtafel | Washtafel | 2000000 |
| PRD - 1196 | Tempat Jas | Tempat jas | 1500000 |
| PRD - 1197 | Bed 200 x 200 | Tempat tidur | 5000000 |
| PRD - 1198 | Lemari 2 pintu | Lemari Baju | 2000000 |
| PRD - 1199 | Lemari 3 pintu | Lemari Baju | 3000000 |
| PRD - 1200 | Lemari 4 pintu | Lemari Baju | 4000000 |
| PRD - 1201 | Lemari 5 pintu | Lemari Baju | 5000000 |
| PRD - 1202 | Meja Rias | Meja Rias | 1500000 |
| PRD - 1203 | Lemari 4 laci | Lemari Baju | 2000000 |
| PRD - 1204 | Lemari 3 laci | Lemari Baju | 1500000 |
| PRD - 1205 | Buffet TV 3 laci | Lemari TV | 2000000 |
| PRD - 1206 | Buffet TV 2 pintu 3 laci | Lemari TV | 1700000 |
| PRD - 1207 | Rak Pejangan | Rak | 2000000 |

Gambar 3.6 Tampilan tb_produk

Gambar 3.6 merupakan tampilan tb_produk yang berisi kode_produk, nama_produk, jenis_produk dan harga_satuan produk dari toko orba express.

```
MariaDB [db_penjualan]> SELECT *FROM tb_persediaan;
```

| kode_produk | jml_terjual |
|-------------|-------------|
| PRD - 1183 | 20 |
| PRD - 1184 | 7 |
| PRD - 1190 | 8 |
| PRD - 1195 | 3 |
| PRD - 1196 | 5 |
| PRD - 1197 | 9 |
| PRD - 1199 | 2 |
| PRD - 1200 | 10 |
| PRD - 1204 | 10 |
| PRD - 1205 | 12 |
| PRD - 1206 | 17 |
| PRD - 1208 | 18 |
| PRD - 1209 | 20 |
| PRD - 1210 | 22 |

Gambar 3.7 Tampilan tb_persediaan

Gambar 3.7 merupakan tampilan tb_persediaan yang berisi kode_produk, dan jml_terjual produk dari toko orba express dan dapat melihat produk yang paling banyak laku terjual.

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan hasil penelitian perancangan data warehouse pada Orba Express Furniture, maka penulis dapat mengambil kesimpulan sebagai berikut.

- 1) Perancangan data warehouse yang telah dibuat dapat membantu pemilik usaha dalam menentukan jumlah produk yang akan di stok pada toko.
- 2) Proses integrasi data warehouse menggunakan konsep ETL.
- 3) Hasil analisis dari perancangan data warehouse menggunakan tabel fakta.

4.2 Saran

Berdasarkan hasil penelitian dan kesimpulan yang didapat, penulis memberikan saran sebagai berikut.

- 1) Menerapkan data warehouse dapat membantu proses pengambilan keputusan dengan cepat dan mudah.
- 2) Data warehouse dapat diterapkan dalam perusahaan besar, usaha kecil, dan segala usaha yang masih di bangun.

REFERENSI

- Dharmayanti, Dian, Adam Mukharil Bachtiar, and Andri Heryandi. 2013. "Pemodelan Data Warehouse." *Jurnal Majalah Ilmiah Unikom* 12(2): 151–68. <https://jurnal.unikom.ac.id/jurnal/pemodelan-data-warehouse-pada.4k>.
- Firman, Astria et al. 2016. "Sistem Informasi Perpustakaan Online Berbasis Web." *E-Journal Teknik Elektro Dan Komputer* 5(2): 29–36.
- Khotimah, Khusnul, and Sriyanto -. 2016. "Perancangan Dan Implementasi Data Warehouse Untuk Mendukung Sistem Akademik (Studi Kasus Pada Stkip Muhammadiyah Kotabumi)." *Jurnal Teknologi Informasi Magister Darmajaya* 2(01): 94–107.
- Pendukung, Sistem et al. 2013. "Desain Dan Implementasi Data Warehouse Untuk." 2(1): 2–5.
- Wibisono, Yudi. 2014. "Pengantar Pentaho Data Integration (Kettle)." 5: 20.