

## **Implementasi Datawarehouse Dalam Menentukan Tabel Fakta Melalui Proses ETL Di Alwy Minimart**

Luthfi Hilmansyah<sup>1</sup>, Siti Maesaroh<sup>2</sup>

<sup>1,2</sup>Teknik Informatika, Sekolah Tinggi Teknologi YBS Internasional  
<sup>1</sup>awesomeluthfi@gmail.com, <sup>2</sup>sitimaesaroh40@gmail.com

### **Abstract**

*In the current information era, data growth occurs very quickly, so data processing is needed that is able to analyze historical data from various aspects. Datawarehouse is one solution that can solve this problem. Datawarehouse is a relational database designed for query and analysis purposes. The datawarehouse contains historical data from transaction data sources, and various other data sources. Datawarehouse can provide dynamic reports because they can be viewed from various dimensions and have the ability to be detailed in more detail according to user needs. However, in the modeling process, it still requires a more specific study in accordance with the problems being handled. Especially related to the specification of the data subject for the management needs of a system which is the basis for decision making, because in each data dimension it is possible to solve in more detail. Based on these problems, this study proposes a data dimension model design to automate business information in the datawarehouse so that it can be used to support business decision-making processes accurately. The approach used is to apply the Star Schema model to quickly provide answers to requests for dimensional analysis processes based on Online Analytical Processing (OLAP). The results of the research are useful for the process of developing more complex facts based on data needs, as well as the dimensions of the data that have been made into a datamart, which can be used in the future.*

**Keywords:** *Datawarehouse, Database, Query, Data Dimension, Star Schema, Online Analytical Processing (OLAP), Datamart*

### **Abstrak**

Pada era informasi sekarang, pertumbuhan data terjadi sangat cepat, sehingga diperlukan pengolahan data yang mampu menganalisa data histori dari berbagai aspek. *Datawarehouse* salah satu solusi yang bisa mengatasi persoalan tersebut. *Datawarehouse* merupakan *database* relasional yang dirancang untuk keperluan *query* dan analisis. *Datawarehouse* berisikan data historis yang berasal dari sumber data transaksi, dan berbagai sumber data lainnya. *Datawarehouse* dapat memberikan laporan bersifat dinamis karena dapat dilihat dari berbagai dimensi dan memiliki kemampuan untuk diperinci lebih detail sesuai kebutuhan pengguna. Namun di dalam proses pemodelannya, masih memerlukan kajian yang lebih spesifik sesuai dengan persoalan yang ditangani. Terutama terkait dengan spesifikasi subjek data untuk kebutuhan pengelolaan suatu sistem yang menjadi dasar dalam pengambilan keputusan, karena pada setiap dimensi data dimungkinkan untuk dilakukan pemecahan lebih detail lagi. Berdasarkan persoalan tersebut, penelitian ini mengusulkan rancangan model dimensi data untuk mengotomatisasi informasi bisnis pada *datawarehouse* sehingga dapat digunakan untuk mendukung proses pengambilan keputusan bisnis secara akurat. Pendekatan yang digunakan menerapkan model *Star Schema* untuk menyajikan jawaban permintaan proses analisis yang bersifat dimensional secara cepat berdasarkan *Online Analytical Processing (OLAP)*. Hasil penelitian bermanfaat untuk proses pengembangan fakta yang lebih kompleks berdasarkan kebutuhan data, serta dimensi data yang telah dibuat menjadi suatu *datamart*, yang dapat dimanfaatkan dikemudian hari.

**Kata kunci:** *Datawarehouse, Database, Query, Data Dimension, Star Schema, Online Analytical Processing (OLAP), Datamart*

### **1. Pendahuluan**

*Minimarket* kini menghadapi dinamika serta daya saing mencari pasar yang lebih baik. Perkembangan dari pada *minimarket* dikarenakan oleh berbagai faktor. Faktor bertambahnya penduduk berpotensi baik kepada *minimarket*. Pendapatan perkapita menawarkan efek pada daya beli masyarakat. Faktor lain semacam kebiasaan masyarakat perkotaan yang mengharapkan kenyamanan, kebersihan, variasi produk, serta lokasi yang dekat dengan daerah tinggal. Lokasi gerai yang

berdekatan menjadi persaingan yang kuat antar *minimarket*. Selain itu, keuntungan yang didapat dari *retail market* sekitar 7% hingga 15% dari omzet [1]. Deskripsi fakta tersebut memiliki konsekuensi terhadap pertumbuhan informasi yang sangat besar, sehingga dibutuhkan suatu sistem yang dapat mengelola data secara menyeluruh. Selain itu, diperlukan juga suatu penyimpanan atau gudang data sebagai salahsatu wadah pemodelan data yang dapat menghasilkan suatu informasi yang bermanfaat untuk kebutuhan eksekutif.

OLAP merupakan suatu metode pendekatan untuk menyajikan jawaban dari permintaan proses analisis yang bersifat dimensional secara cepat, yaitu desain dari aplikasi dan teknologi yang dapat mengoleksi, menyimpan, memanipulasi suatu data multidimensi untuk tujuan analisis [2].

OLAP digunakan untuk keperluan analisis data oleh manajer, eksekutif bisnis dan analis pemasaran. *Datawarehouse* dan OLAP telah muncul sebagai teknologi yang terkemuka yang dapat memfasilitasi penyimpanan data, pengaturan dan pencarian data. Teknologi OLAP memungkinkan *datawarehouse* untuk digunakan secara efektif untuk kepentingan analisis *online*, menyediakan respon cepat terhadap *iterative complex analytical queries*. Mutidimensional data model OLAP dan teknik *data aggregation* mengatur dan menyederhanakan data dalam jumlah yang besar sehingga dapat dievaluasi secara cepat menggunakan *online analysis* dan *graphical tools*.

*Datawarehouse* merupakan suatu sistem yang mengambil serta mengkonsolidasikan data dengan cara periodik dari suatu sumber data ke suatu kawasan penyimpanan data yang bersifat dimensional maupun relasional. Suatu *datawarehouse* merupakan suatu koleksi data yang bisa digunakan untuk menunjang pengambilan keputusan manajemen, yang berorientasi subjek (topik), terpadu, *time variant*, dan tidak mudah berubah [3].

Pemodelan *datawarehouse* yang paling umum adalah model dimensi data, yang dilakukan dengan berbagai macam model. Salah satunya adalah model *Star Schema*, model ini digunakan oleh Amborowati (2009) dalam penelitiannya yang merekomendasikan perlunya pengembangan subjek data agar lebih sesuai dengan kebutuhan pengelolaan suatu sistem [4]. Dimensi data inilah yang menjadi subjek informasi untuk menjadi bahan dalam pengambilan keputusan, karena pada setiap dimensi data dimungkinkan untuk dilakukan pemecahan lebih detail lagi. Dengan demikian sumber data yang bisa diolah untuk menjadi

informasi bisa menjadi lebih banyak dan detail [5].

Berdasarkan permasalahan tersebut, pembuatan *datawarehouse* untuk kebutuhan proses transaksi merupakan salah satu bentuk pemecahan masalah yang muncul. *Datawarehouse* ini menyediakan sudut pandang dari berbagai perspektif analisis bisnis dan pembuatan keputusan sehingga dibutuhkan untuk pemrosesan transaksi dan pemecahan masalah laporan penjualan yang lebih efektif. Dalam penelitian ini, fokus yang menjadi kajian utama adalah bagaimana mengembangkan suatu desain *datawarehouse* dengan menerapkan metode OLAP.

## 2. Kajian pustaka

### 2.1 Teori Terkait

#### Datawarehouse

*Datawarehouse* adalah *database* relasional yang didesain untuk keperluan *query* serta analisis. *Datawarehouse* berisikan data historis yang bersumber dari sumber data transaksi, tetapi juga bersumber dari beberapa sumber data lainnya. *Datawarehouse* memisahkan beban kerja analisis dari transaksi dan memungkinkan perusahaan untuk menggabungkan data dari beberapa sumber [6].

#### Karakteristik *Datawarehouse*

Menurut Bapak *Datawarehouse* W. H. Inmon karakteristik *datawarehouse* adalah sebagai berikut:

a) *Subject-oriented*

*Datawarehouse* diorganisasikan subjek utama perusahaan (seperti pelanggan, produk, dan penjualan) daripada area utama aplikasi (seperti pelanggan, faktur, pengendalian stok, dan penjualan produk).

b) *Integrated*

Karena sumber data datang secara bersamaan dari berbagai *enterprise-wide application system*.

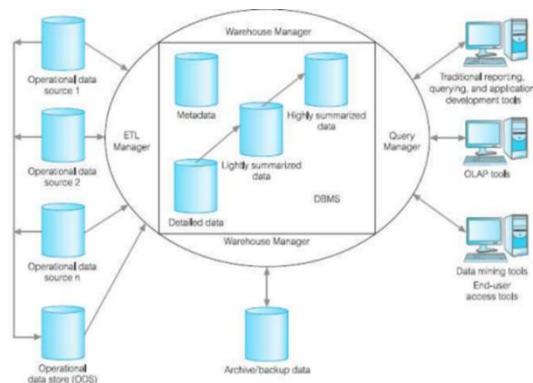
c) *Time variant*

Karena data dalam *warehouse* hanya akurat dan berlaku di beberapa titik di suatu waktu atau selama beberapa interval waktu.

d) *Non-volatile*

Data tidak di perbarui secara *real time* tapi memperbarui dari operasional sistem secara teratur. Data baru selalu ditambahkan sebagai suplemen ke *database*, *database* terus menyerap data baru secara bertahap dengan mengintegrasikan data sebelumnya.

### Arsitektur Datawarehouse



**Gambar 1.** Arsitektur Datawarehouse Menurut Connolly & Begg

dalam membangun sebuah *datawarehouse* menurut Girsang (2017) yang dikutip dari Ralph Kimball, dikenal dengan *nine-step design methodology*, jika langkah-langkah dalam *nine-step design methodology* dilakukan secara sistematis, maka dapat membangun sebuah *datawarehouse* yang baik.

a) *Choose the Process*

Memilih proses berarti menentukan subjek utama.

b) *Choose the Grain*

Memilih grain berarti menentukan apa yang akan diwakili atau dipresentasikan oleh sebuah tabel fakta.

c) *Identify and Conform the Dimensions*

Mengidentifikasi dan menghubungkan tabel dimensi dengan tabel fakta.

d) *Choose the Facts*

Grain dari suatu tabel fakta menentukan fakta-fakta yang bisa digunakan.

e) *Store pre-Calculations in the Fact Table*

Pada tahap ini, hasil perhitungan pada suatu atribut perlu dipertimbangkan untuk disimpan dalam database.

f) *Round Out the Dimension Tables*

Dari dimensi-dimensi yang telah diidentifikasi, dibuat deskripsi yang memuat informasi terstruktur mengenai atribut-atribut pada tabel dimensi. Tabel dimensi tersebut harus diberi keterangan secara lengkap dan mudah dipahami oleh pengguna.

g) *Choose the Durations of the Database*  
Durasi waktu dari data-data yang akan dimasukkan ke dalam datawarehouse akan ditentukan pada tahap ini.

h) *Determine the Need to Track Slowly Changing Dimension*

Dimensi dapat berubah dengan lambat dan menjadi sebuah masalah.

i) *Decide Physical Design*

Pada tahap ini, dilakukan perancangan fisik dari datawarehouse.

### Online Analytical Processing (OLAP)

*Datawarehouse* dan OLAP telah muncul sebagai teknologi yang terkemuka yang dapat memfasilitasi penyimpanan data, pengaturan dan pencarian data. Teknologi OLAP memungkinkan *datawarehouse* untuk digunakan secara efektif untuk kepentingan analisis online, menyediakan respon cepat terhadap *iterative complex analytical queries*. Mutidimensional data model OLAP dan teknik *data aggregation* mengatur dan menyederhanakan data dalam jumlah yang besar sehingga dapat dievaluasi secara cepat menggunakan *online analysis* dan *graphical tools*. OLAP merupakan suatu alat yang disediakan oleh *datawarehouse* untuk kepentingan pengambilan keputusan yang intensif dimana data disimpan kedalam model multidimensional. OLAP biasanya digunakan untuk mendeskripsikan analisis dari informasi kompleks *datawarehouse* [7].

### Extract, Transform dan Load (ETL)

a) *Extract*

Ekstrak merupakan proses pemilahan dan pengambilan data dari satu atau lebih sumber data. Pada proses ekstrak, pertama-tama data diambil dari *database* operasional. Proses ekstraksi informasi dari sumber asli dan dipindahkan ke dalam *datawarehouse* merupakan hal yang sangat penting. Hal ini disebabkan

sejumlah besar volume data akan dipindahkan dari sistem sumber (*legacy*) ke dalam *datawarehouse* dengan ekstraksi data dilakukan secara otomatis. Hanya item yang dibutuhkan yang dipindahkan ke dalam *datawarehouse*. Informasi yang masuk ke dalam *datawarehouse* dapat datang dari internal sistem (*legacy system*) maupun dari eksternal organisasi [8].

b) *Transform*

Transformasi adalah serangkaian aturan atau fungsi untuk *extract* data, yang mana menentukan bagaimana data akan digunakan untuk analisis dan dapat melibatkan data transformasi seperti pengkodean data, penggabungan data, pemisahan data, perhitungan data, dan pembuatan *Surrogate Key*. *Output* dari transformasi adalah data yang bersih dan konsisten dengan data yang sudah di *warehouse* dan selanjutnya adalah dalam bentuk yang siap untuk dianalisis oleh pengguna *warehouse* [9].

c) *Loading*

Fase *load* merupakan tahapan yang berfungsi untuk memasukkan data ke dalam target akhir, yang biasanya ke dalam suatu *datawarehouse*. Jangka waktu proses ini tergantung pada kebutuhan organisasi. Beberapa *datawarehouse* dapat setiap minggu mengisi keseluruhan informasi yang ada secara kumulatif, data diubah, sementara *datawarehouse* yang lain (atau bagian lain dari *datawarehouse* yang sama) dapat menambahkan data baru dalam suatu bentuk yang historikal, contohnya setiap jam [10].

### **Multidimensional Data Model**

Multidimensional data model merupakan teknik desain logis yang bertujuan untuk menyajikan data secara standar, merupakan bentuk intuitif yang memungkinkan adanya akses performa yang tinggi. Multidimensional data model juga diketahui sebagai data cube yang mana memungkinkan data untuk dimodelkan dan dilihat dari berbagai dimensi atau sudut

pandang. Multidimensional data model mengorganisir dan meringkas data yang berjumlah banyak agar data dapat dievaluasi secara cepat menggunakan analisis online dan graphical tools.

a) *Tabel Dimensi*

Dimensi berisikan referensi informasi mengenai fakta seperti tanggal, produk atau pelanggan. Tabel dimensi berisi informasi tekstual deskriptif dan dimensi atribut digunakan sebagai batasan dalam *query datawarehouse*. Terlebih lagi, tabel dimensi adalah tabel yang diperlukan untuk melengkapi tabel fakta. Tabel dimensi berisi konteks tekstual data yang terkait dengan bisnis proses.

b) *Tabel Fakta*

Tabel fakta adalah tabel besar didalam susunan *datawarehouse* yang menyimpan pengukuran bisnis. Tabel Fakta merepresentasikan data, biasanya dalam bentuk angka dan data tambahan, yang dapat dianalisa dan diperiksa. Contohnya termasuk penjualan, biaya dan, keuntungan.

c) *Surrogate Key*

*Surrogate key* merupakan kunci primer yang bersifat unik untuk mengidentifikasi setiap *record* yang disebut dengan kunci dimensi *surrogate*. Kunci ini merupakan sebuah *identifier* yang dibangkitkan oleh sistem *database* ataupun oleh aplikasi proses ETL. Terlebih lagi *surrogate key* menghindari *datawarehouse* dari adanya perubahan data oleh *user*.

d) *Star Schema*

*Star schema* merupakan dimensional data model yang memiliki tabel fakta ditengahnya dan dikelilingi oleh tabel dimensi yang di denormalisasi. *Star schema* ditandai oleh satu atau lebih tabel fakta yang sangat besar yang berisikan informasi mutlak pada *datawarehouse* serta sejumlah tabel dimensi yang lebih kecil, yang masing-masing berisikan informasi tentang entri untuk atribut tertentu pada tabel fakta.

## Pentaho Data Integration (PDI)

Pentaho adalah kumpulan aplikasi business intelligence yang bersifat free open source software (FOSS) dan berjalan diatas platform java. Selain sifatnya gratis dalam bentuk lisensi, pentaho juga dapat dibeli dalam bentuk service level agreement (SLA) dan dipaketkan dalam versi enterprise edition yang sifatnya annual subscription atau perlu langganan.

Pentaho Data Integration atau Kettle adalah tools yang memiliki kemampuan extract, transform, dan load (ETL) pada multi platform database. Script dari desain dapat disimpan dalam bentuk file ataupun repository. Selain itu, pada tools ini terdapat cukup banyak 'steps' untuk mengatur workflow control (JOB), dan data workflow (Transformation).

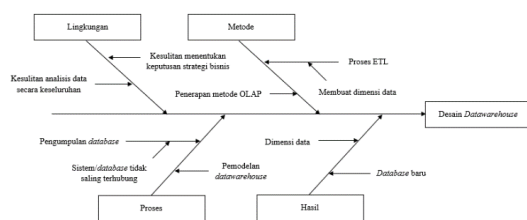
### 2.2 Pekerjaan Terkait

Pada salah satu penelitian terkait dan yang hampir mirip dengan penelitian yang sedang dilakukan, yaitu penelitian dengan judul Perancangan Datawarehouse pada Perpustakaan Yayasan Lentera Insan, membahas masalah data perlu diolah terlebih dahulu menjadi sebuah informasi. Proses pengolahan data dapat dilakukan diberbagai tempat, misalkan di database operasional, aplikasi operasional, maupun teknologi datawarehouse. Dengan metode yang digunakan adalah metode grounded (grounded research) yaitu suatu metode penelitian berdasarkan pada fakta dan menggunakan analisis perbandingan dengan tujuan mengadakan generalisasi empiris, menetapkan konsep, membuktikan teori, mengembangkan teori, pengumpulan dan analisis data dalam waktu yang bersamaan.

## 3. Metode Penelitian

### 3.1 Fishbone Diagram

Bentuk *Fishbone Diagram* dari penelitian ini dikembangkan untuk membuat rancangan model dimensi data..



**Gambar 2** Fishbone Diagram

### 3.1.1 Lingkungan

a) Kesulitan analisis data secara keseluruhan

Proses ini merupakan permasalahan utama dibuatnya penelitian mengenai desain datawarehouse, karena aplikasi yang sedang berjalan digunakan di tiap cabang bisnis memiliki sistem penjualan dan database yang terpisah. Adanya datawarehouse memudahkan dalam melakukan proses analisis data untuk penentuan keputusan.

b) Kesulitan menentukan keputusan strategi bisnis

Proses bisnis yang terjadi adalah bisnis dengan memiliki beberapa cabang, hal ini mengharuskan pemilik bisnis dapat melakukan analisis terhadap proses penjualan secara keseluruhan, yang kemudian menjadi permasalahan ketika pimpinan/manajer perusahaan ingin menentukan strategi penjualan atau pengambilan keputusan strategis.

### 3.1.2 Proses

a) Pengumpulan database

Proses pengumpulan database ini adalah suatu proses penarikan database dari seluruh cabang usaha sebagai tahap awal dalam menentukan skema datawarehouse yang akan dibuat. Pengumpulan database ini dilakukan karena tidak saling terhubungnya sistem yang digunakan di masing-masing cabang.

b) Pemodelan datawarehouse

Proses pemodelan datawarehouse adalah langkah awal membuat skema pembuatan dimensi database baru dengan menggunakan *star schema* sebagaimana pada bab sebelumnya telah di bahas.

### 3.1.3 Metode

a) Penerapan OLAP

OLAP dibangun dengan tujuan agar pengguna dapat melakukan analisis multidimensi sehingga menghasilkan informasi yang diperlukan.

b) Proses ETL

Proses pemodelan data ini bertujuan adalah langkah untuk membuat skema

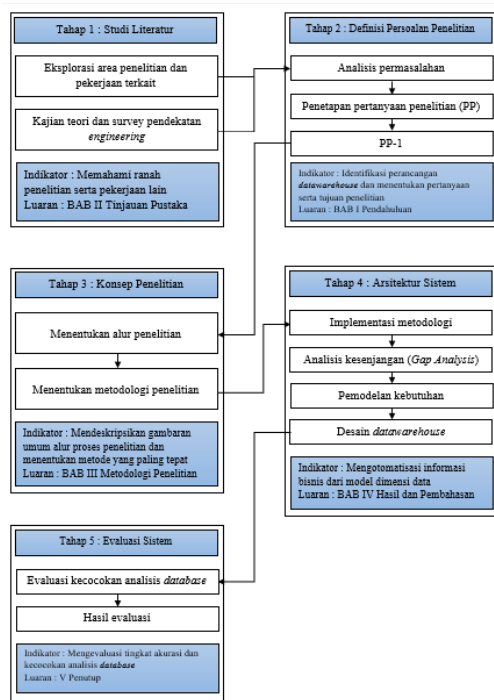
dimensi data sehingga menghasilkan database baru dengan berbagai dimensi data baru yang dapat dijadikan sebagai olah data dalam pengambilan keputusan.

### 3.1.4 Hasil

- a) Dimensi data  
Dimensi data ini dihasilkan dari proses ETL yang menghasilkan skema database.
- b) Database baru  
Database baru ini yang dapat digunakan ke tahap visualisasi untuk penentuan keputusan dengan memanfaatkan dimensi data.

## 3.2 Tahapan Penelitian

Adapun tahapan penelitian ini dapat dilihat pada gambar berikut.



Gambar 3 Tahapan Penelitian

### 3.2.1 Studi Literatur

Indikator pada tahap ini yaitu untuk memahami domain penelitian dan pekerjaan terkait. Aktivitas yang dilakukan adalah mengeksplorasi penelitian dan pekerjaan terkait, diantaranya melakukan kajian teori tentang datawarehouse dan melakukan survey pendekatan engineering.

### 3.2.2 Definisi Persoalan Penelitian

Terdapat 3 hal yang dilakukan pada tahap ini.

- a) Analisis Permasalahan  
Penelitian diawali dengan melakukan wawancara untuk identifikasi

permasalahan yang dialami, khususnya terkait perancangan *datawarehouse*.

- b) Analisis Kebutuhan

Tahap ini juga dilakukan melalui wawancara untuk identifikasi kebutuhan agar mampu memberikan solusi terhadap permasalahan serta melakukan identifikasi terhadap kebutuhan perancangan *datawarehouse*.

- c) Analisis Data

Sebelum melakukan analisis data, dilakukan pengambilan sampel data omzet penjualan. Data tersebut kemudian dianalisis agar dapat mempermudah perancangan dan pemodelan *datawarehouse*.

### 3.2.3 Konsep Penelitian

- a) Menentukan alur penelitian

Menjelaskan tentang bagaimana alur proses penelitian yang sedang diteliti. Tahapan ini adalah upaya-upaya untuk mencari jawaban dari suatu masalah dengan cara pengumpulan, pengolahan, penyajian dan analisis data secara sistematis.

- b) Menentukan metodologi penelitian

Tahap penentuan metode yang akan digunakan untuk menyelesaikan permasalahan. Dalam penelitian ini menggunakan metode OLAP melalui proses ETL.

### 3.2.4 Arsitektur Sistem

- a) Implementasi Metodologi

- 1) Proses *Extraction, Transformation* dan *Loading* (ETL).
- 2) Penerapan OLAP, dibangun dengan tujuan agar pengguna dapat melakukan analisis multidimensi sehingga menghasilkan informasi yang diperlukan untuk mendukung keputusan.

- b) Analisis Kesenjangan (Gap Analysis)

Deskripsi penelitian terkait yang terdapat dari rangkuman beberapa paper penelitian sebelumnya, deskripsi tersebut berisi tentang kelemahan atau kesenjangan masalah yang belum terpecahkan oleh peneliti sebelumnya.

Deskripsi tersebut menegaskan perlunya penelitian lebih lanjut berdasarkan kekurangan penelitian-penelitian sebelumnya.

- c) Pemodelan kebutuhan
  - 1) Perancangan datawarehouse, Perancangan datawarehouse menggunakan nine step methodology yang terdiri dari 9 tahap.
  - 2) Skema datawarehouse, Pemodelan datawarehouse yang paling umum adalah model dimensi data, salah satunya adalah Star Schema, dengan menentukan data apa yang ingin dilihat oleh pengguna (besarannya) dan bagaimana pengguna melihat data tersebut (kondisi atau dimensinya).

### 3.2.5 Evaluasi Sistem

Pengembang sistem melakukan uji coba sebelum aplikasi OLAP disampaikan kepada pengguna dengan cara membuat laporan sesuai dengan kebutuhan. Namun jika belum sesuai, maka kembali ke tahap desain untuk memperbaiki rancangan maupun pemodelan.

## 4. Hasil dan Pembahasan

### 4.1 Perancangan (*Design*)

Proses design akan menerjemahkan syarat kebutuhan kedalam sebuah desain tabel suatu database yang akan dilakukan proses ETL.

#### 4.1.1 Desain Database

Sebelum melakukan proses eksekusi database, hal pertama yang disiapkan adalah dua buah database yaitu.

- a) Database phi-minimart (nama database yang akan dijadikan studi kasus)
- b) Database dw-phi (nama database baru sebagai database penyimpanan hasil proses ekstraksi)

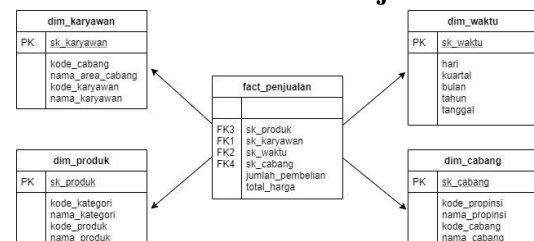
#### 4.1.2 Memahami Struktur Tabel Database Masukan

Tabel tersebut mendeskripsikan detail nama tabel yang ada pada database phi-minimart.sql, yang menjadi database minimart, yang akan menjadi sumber database untuk kemudian dikembangkan menjadi tabel dimensi dan tabel fakta.

**Tabel 1.** Struktur *database* phi-minimart.sql  
*Database: phi-minimart.sql*

Nama Tabel	Nama Field
<b>ms_karyawan</b>	kode_cabang : varchar (10)
	kode_karyawan : varchar (10)
	nama_depan : varchar (8)
	nama_belakang : varchar (9)
<b>ms_cabang</b>	jenis_kelamin : varchar (1)
	kode_cabang : varchar (10)
<b>ms_produk</b>	nama_cabang : varchar (100)
	kode_kota : varchar (8)
	kode_item : varchar (7)
	kode_produk : varchar (12)
<b>ms_kategori</b>	kode_kategori : varchar (7)
	nama_produk : varchar (100)
	unit : int (11)
	kode_satuan : varchar (4)
<b>ms_kategori</b>	kode_kategori : varchar (7)
	nama_kategori : varchar (17)
<b>ms_kota</b>	kode_kota : varchar (8)
	nama_kota : varchar (16)
	kode_propinsi : varchar (3)
<b>ms_propinsi</b>	kode_propinsi : varchar (3)
	nama_propinsi : varchar (25)
<b>ms_harga_harian</b>	kode_produk : varchar (12)
	tgl_berlaku : datetime
	kode_cabang : varchar (10)
	harga_berlaku_cabang : int (11)
	modal_cabang : int (11)
	biaya_cabang : int (11)
<b>tr_penjualan</b>	tgl_transaksi : datetime
	kode_cabang : varchar (10)
	kode_kasir : varchar (10)
	kode_item : varchar (7)
	kode_produk : varchar (12)
	jumlah_pembelian : int (11)

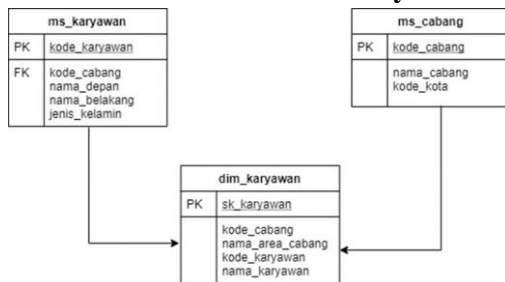
#### 4.1.3 Desain Tabel Fakta Penjualan



**Gambar 4** Desain tabel fakta penjualan  
 Gambar di atas merupakan desain star schema dari database yang akan dilakukan proses ekstraksi sehingga didapatkan tabel baru berupa

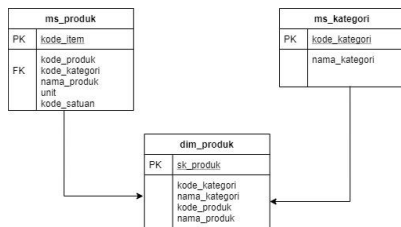
dim\_karyawan, dim\_produk, dim\_waktu, dim\_cabang serta fact\_penjualan.

#### 4.1.4 Desain Tabel Dimensi Karyawan



**Gambar 5** Desain tabel dimensi karyawan  
 Tabel dimensi karyawan ini didapatkan dari hasil gabungan tabel ms\_karyawan dan ms\_cabang pada database phi\_minimart. Hal ini bertujuan untuk mendapatkan data baru berupa “Masing-masing karyawan berada di cabang yang mana”.

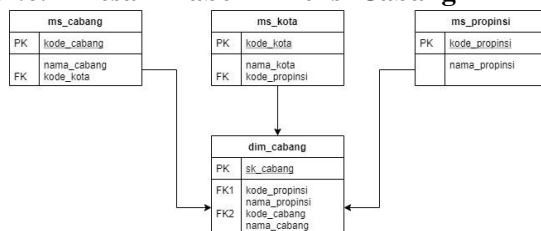
#### 4.1.5. Desain Tabel Dimensi Produk



**Gambar 6** Desain tabel dimensi produk

Tabel dimensi produk ini didapatkan dari hasil gabungan tabel ms\_produk dan ms\_kategori pada database phi\_minimart. Hal ini bertujuan untuk mendapatkan data baru berupa “Klasifikasi produk berdasarkan kategori produk”.

#### 4.1.6. Desain Tabel Dimensi Cabang



**Gambar 7** Desain tabel dimensi cabang  
 Tabel dimensi cabang ini didapatkan dari hasil gabungan tabel ms\_cabang, ms\_kota dan ms\_propinsi pada database phi\_minimart. Hal ini bertujuan untuk mendapatkan data dimensi baru berupa “Klasifikasi cabang berdasarkan kota”. Namun, pada bagian dimensi cabang ini memiliki 3 masukan yaitu ms\_cabang, ms\_kota dan ms\_propinsi, karena pada database phi\_minimart tabel kota dan propinsi dibuat

terpisah sehingga perlu dua kali proses penggabungan, yang pertama adalah penggabungan ms\_cabang dengan ms\_kota, hasilnya digabungkan dengan ms\_propinsi. Berikut desain untuk proses pengambilan data untuk dimensi cabang.

#### 4.1.7. Desain Tabel Dimensi Waktu



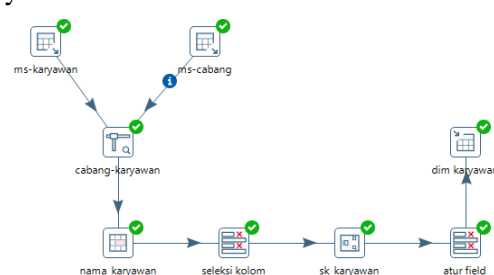
**Gambar 8** Desain tabel dimensi waktu  
 Design awal yang dilakukan membuat tabel dimensi waktu dengan nama output tabel dim\_waktu pada database dw-phi. Sebelum melakukan desain tabel waktu, sebelumnya harus mendefinisikan terlebih dahulu tanggal awal dan hari selanjutnya, karena tabel waktu ini dibuat dari membuat field di pentahonya.

#### 4.2 Implementasi

Tahap implementasi meliputi proses Extract, Transform dan Load (ETL) dengan menggunakan tools yang bernama Pentaho Data Integration.

##### 4.2.1 Membuat Tabel Dimensi Karyawan

Tabel dimensi karyawan didapatkan dari hasil penggabungan dua tabel yang ada di database phi\_minimart, yaitu tabel ms\_karyawan dan ms\_cabang, sehingga dibutuhkan dua step pada pentaho yaitu step Table Input. Untuk menghasilkan tabel dimensi karyawan diperlukan proses ETL pada database yang akan diolah. Berikut desain ETL untuk tabel dimensi karyawan.



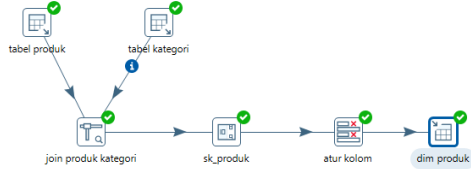
**Gambar 9** Desain ETL tabel dimensi karyawan

##### 4.2.2 Membuat Tabel Dimensi Produk

Tabel dimensi produk difungsikan untuk mendapatkan data produk berdasarkan kategori. Tabel dimensi produk didapatkan dari



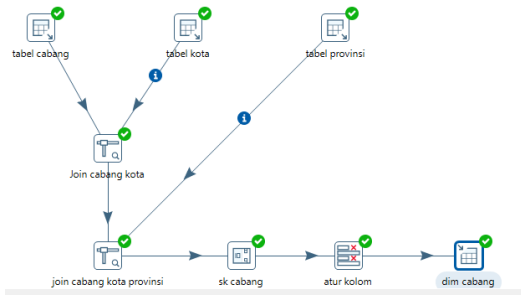
penggabungan tabel ms\_kategori dan ms\_produk pada database phi\_minimart, dengan catatan koneksi dengan database sumber telah dilakukan. Berikut beberapa proses yang dilakukan untuk mendapatkan tabel dimensi produk pada database dw-phi.



**Gambar 10** Desain ETL tabel dimensi produk

#### 4.2.3 Membuat Tabel Dimensi Cabang

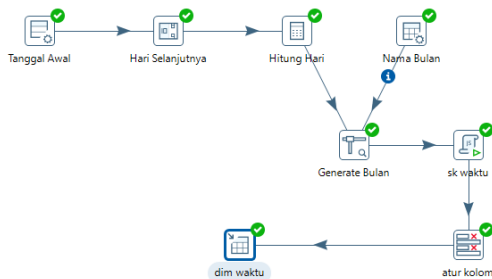
Tabel dimensi cabang difungsikan untuk mendapatkan data baru berupa klasifikasi data cabang berdasarkan kota dan propinsi. Dibutuhkan tiga inputan dalam proses ini, karena pada studi kasus ini tabel kota dan tabel propinsi dibuat secara terpisah, maka secara logika perlu penggabungan diantara keduanya.



**Gambar 11** Desain ETL tabel dimensi cabang

#### 4.2.4 Membuat Tabel Dimensi Waktu

Tahap pembuatan dimensi waktu ini berbeda dengan pengolahan data pada dimensi yang lainnya, pada dimensi waktu ini data yang ingin dicapai adalah data hari, tanggal, tahun, kuartal.

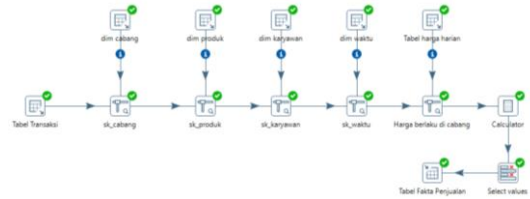


**Gambar 12** Desain ETL tabel dimensi waktu

#### 4.2.5 Membuat Tabel Fakta Penjualan

Proses pembuatan tabel fakta ini merupakan penggabungan dari beberapa tabel dimensi yang telah dibuat sebelumnya, berikut desain tabel

fakta penjualan. Tabel fakta penjualan didapatkan dengan mengolah sumber tabel tr\_penjualan pada database phi\_minimart dan semua tabel dimensi pada database dw-phi.



**Gambar 13** Desain ETL tabel fakta penjualan

**4.2.6 Hasil Tabel Fakta Penjualan**  
 Tabel fakta penjualan didapatkan dari proses pengolahan database phi-minimart melalui proses ETL sehingga menghasilkan tabel fakta sebagai berikut.

**Gambar 14** Hasil tabel fakta penjualan

## 5. Simpulan

Berdasarkan seluruh hasil pembahasan penelitian yang telah dilakukan, menghasilkan beberapa simpulan sebagai berikut:

- 1) Penelitian ini telah berhasil merancang model dimensi data berbasis model star schema untuk menunjang pengambilan keputusan datawarehouse penjualan.
- 2) Penelitian ini menghasilkan suatu analisis data yang diperlukan untuk membuat sebuah datawarehouse dalam perusahaan minimarket yang memiliki cabang atau waralaba dalam pelaksanaan proses bisnisnya, sehingga dihasilkan sebuah data berupa fakta yang dapat digunakan oleh eksekutif.
- 3) Penelitian ini menghasilkan database baru berupa tabel-tabel dimensi dan tabel fakta, yang menjadi output penelitian. Selain itu, database tersebut dapat dikembangkan lebih kompleks sesuai dengan kebutuhan eksekutif.

- 4) Penelitian ini memberikan skema proses Extract, Transform dan Load pada suatu database.
- 5) Model atau tipe keputusan yang dapat didukung dalam proses pengambilan keputusan datawarehouse penjualan diantaranya adalah mengetahui data karyawan berdasarkan cabang, klasifikasi produk berdasarkan kategori dan mengetahui fakta penjualan produk dari seluruh cabang minimarket yang ada.
- 6) Hasil penelitian ini bermanfaat untuk proses pengembangan fakta lain yang lebih kompleks berdasarkan kebutuhan data yang diinginkan, serta dimensi data yang telah dibuat menjadi suatu datamart yang dapat dimanfaatkan dikemudian hari. Selain itu, eksekutif (pemilik usaha/manajemen) dapat dengan mudah melihat data penjualan dari keseluruhan cabang yang ada secara ringkas.

Berdasarkan hasil penititan dari simpulan dan analisis yang telah dilakukan pada datawarehouse minimarket tersebut masih dapat dikembangkan lebih jauh dan lebih kompleks serta dengan skala data yang lebih besar.

#### Daftar Pustaka

- [1] Nurzani, Z. (2019). IMPLEMENTASI ASSOCIATION RULE DALAM MENGANALISISTRANSAKSI PENJUALAN 212 MART KUTO PALEMBANG . *Jurnal Sistem Informasi, Universitas Sriwijaya*.
- [2] Sinaga, A. S. (2019). Implementasi OLAP Menggunakan Dashboard Holistic Software Pada LPPM STMIK Pelita Nusantara. *Jurnal Penelitian Teknik Informatika Universitas Prima Indonesia (UNPRI) Medan, Volume 2 Nomor 1, e-ISSN : 2541-2019, 55*.
- [3] Asahar Johar, A. V. (2015). Aplikasi Business Intelegence (BI) Data Pasien Rumah Sakit M.Yunus Bengkulu dengan Menggunakan Metode OLAP (Online Analytical Processing). *Jurnal Rekursif, Vol. 3 No. 1, ISSN 2303-0755, 14*.
- [4] Akhmad Dahlan, dkk. (2013). Perancangan Data Warehouse Perpustakaan Perguruan Tinggi XYZ Menggunakan Metode Snowflake Schema. *Jurnal Teknologi Informasi Vol. VIII Nomor 24, ISSN : 1907-2430*.
- [5] Syah, R. (2014). Rancang Bangun Data Warehouse Untuk Analisis Strategi Produksi Penjualan Usulan : PT. XYZ. *Jurnal Penelitian Teknik Informatika, TECHSI Vol.4 No.1, 2*.
- [6] Padeli, Allam, M. D., & Suharto, N. R. (2017). PERANCANGAN SISTEM DATA WAREHOUSE PEMBELAJARAN PADA PONDOK PESANTREN AL-FURQON. *CERITA Vol.3 No.1 ISSN 2461-1417*.
- [7] Wijaya, A.F., Sugiarto, A.T. (2017). Proses Extraction, Transformation, and Loading pada Pemodelan Datawarehouse PO. Sumber Alam Kutoarjo. *JUTEI, Volume 1 No.1, ISSN 2579-3675*.
- [8] Safril, A. (2009). PERANCANGAN DAN IMPLEMENTASI DATA WAREHOUSE METEOROLOGI, KLIMATOLOGI, GEOFISIKA DAN BENCANA ALAM. *JURNAL METEOROLOGI DAN GEOFISIKA, VOLUME 10 NOMOR 2, ISSN 1411-3082*.
- [9] Connolly, T. M., & Begg, C. E. (2015). Database systems: a practical approach to design, implementation, and management, Global ed. *Harlow, Pearson Education*.
- [10] Darudiato, S. (2010). PERANCANGAN DATAWAREHOUSE PENJUALAN UNTUK Mendukung Kebutuhan Informasi Eksekutif CEMERLANG SKIN CARE. *Seminar Nasional Informatika 2010 (semnasIF 2010), UPN "Veteran" Yogyakarta, ISSN 1979-2328*.