

PERANCANGAN DATA WAREHOUSE PERPUSTAKAAN PERGURUAN TINGGI XYZ MENGGUNAKAN METODE SNOWFLAKE SCHEMA

Akhmad Dahlan ¹⁾, Ema Utami ²⁾, Emha Taufiq Luthfi ³⁾

¹⁾ Mahasiswa Magister Teknik Informatika Program Pasca Sarjana STMIK Amikom Yogyakarta

^{2,3)} Dosen Magister Teknik Informatika Program Pasca Sarjana STMIK Amikom Yogyakarta
Jl. Ringroad Utara Condongcatur, Depok, Sleman, Yogyakarta 55283

ABSTRAKSI

Keberadaan Perpustakaan sebagai sebuah unit pelayanan teknis dalam sebuah perguruan tinggi adalah sangat penting untuk memberikan layanan terhadap civitas akademika. Dengan semakin banyaknya koleksi pustaka maka sistem basis data yang dibangun harus bisa meningkatkan pelayanan yang berorientasi pada penyediaan data warehouse. Terlebih pada tingkatan manajerial memerlukan informasi yang lengkap, cepat dan akurat untuk mendukung proses dan kegiatan perencanaan, evaluasi, dan pengambilan keputusan yang tepat.

Perancangan data warehouse ditentukan oleh adanya deskripsi kebutuhan informasi yang tepat, pemilihan sumber data yang valid, perancangan data warehouse menggunakan perancangan snowflake schema, dan proses ETL untuk mengintegrasikan, mengekstraksi, membersihkan, mentransformasi serta mempopulasikannya ke dalam data warehouse. Perancangan dengan metode Snowflake schema digunakan karena untuk mengakomodasi lebih banyak tabel dimensi, karena akan memiliki sub-tabel dimensi lain sehingga bisa menghasilkan lebih banyak informasi yang akan dijadikan sebagai bahan untuk mengambil keputusan.

Key word : *data warehouse, perpustakaan, snowflake schema, ETL.*

1. PENDAHULUAN

1.1. Latar Belakang Masalah

Pemanfaatan teknologi *data warehouse* hampir dibutuhkan oleh semua organisasi, tidak terkecuali di Perpustakaan Perguruan Tinggi (PPT). PPT diharapkan mempunyai *data warehouse*, agar proses integrasi data bisa dilakukan dengan mudah. Pada dasarnya proses yang terjadi adalah aktifitas meringkas data yang telah tersimpan di *data warehouse* atau sistem basisdata terkait.

Seorang penyelia perlu mendapatkan data ringkasan data transaksi yang terjadi pada bagiannya atau yang dikenal dengan *transactional information systems* (TPS). Pada tingkat manajemen (*management information systems*), biasanya terjadi peringkasan lebih lanjut terhadap informasi yang dihasilkan oleh penyelia bagian sehingga sesuai dengan kebutuhan para manajer (Amborowati, 2009). Pada tingkat lebih tinggi lagi, informasi harus disajikan sedemikian rupa sehingga dapat mendukung proses pengambilan keputusan. Sistem pada tahap ini dikenal dengan nama *decision support system*. Informasi tertinggi perlu tersedia untuk kebutuhan manajer

perpustakaan dan ringkasan informasi yang jauh lebih detail serta informasinya disediakan dalam sebuah mekanisme tertentu dinamakan *executive information system* (Yuadi, 2007).

Pada tingkatan operasional pelayanan, untuk menangani proses transaksi menggunakan system basis data transaksional. Di dalam basis data tersebut akan menyimpan semua informasi secara detail untuk semua kejadian maupun transaksi yang dilakukan. Data yang tersimpan akan selalu dinamis, sesuai dengan peningkatan terjadinya transaksi dalam basis data tersebut. Akan tetapi, pada tingkatan manajerial informasi yang dibutuhkan pada umumnya adalah merupakan hasil agregasi atau summary dari basis data transaksional terbut. Untuk bisa menyajikan data yang dibutuhkan oleh tingkatan manajerial, keberadaan *data warehouse* sangatlah penting (Prasetyo, 2011). Indikasi pentingnya *data warehouse* tercermin seperti yang diungkapkan oleh Mandias (2012). Dalam penelitiannya menyimpulkan pentingnya *data warehouse* untuk pengolahan data mahasiswa di Universitas Klabat Manado. Proses analisis terhadap data akan bisa dilakukan dengan lebih cepat (Mandias, 2012).

Hal yang paling penting dalam proses pembangunan *data warehouse* adalah proses *modelling* (Patel, 2012). Karena dari model data yang dirancang ini akan bisa diukur sejauh mana *data warehouse* yang dibangun mampu menghasilkan informasi manajerial yang butuhkan. Pemodelan *data warehouse* yang paling umum adalah model dimensi data. Pemodelan *data warehouse* dapat dilakukan dengan berbagai macam metode. Salah satunya adalah model *star schema* yang digunakan oleh Amborowati (2009) dalam melakukan penelitiannya yang merekomendasikan perlunya pengembangan subyek data agar lebih sesuai dengan kebutuhan manajemen perpustakaan.

Salah satu metode lain untuk melakukan pemodelan data warehouse adalah *snowflake schema*. Metode *snowflake schema* adalah pengembangan dari *star schema*, yaitu metode normalisasi untuk tabel dimensi yang ada dalam *star schema* (Ponniah, 2001). Dalam *star schema*, sebuah tabel dimensi tidak memiliki sub-tabel (sub-dimensi) sehingga data yang dapat dijadikan sebagai bahan untuk menghasilkan informasi tidak banyak. Dalam *snowflake schema*, setiap tabel dimensi dapat memiliki sub-tabel dimensi lagi. Hal ini bertujuan untuk meminimalkan data yang berlebihan (*redudancy data*) (Rainardi, 2008). 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.

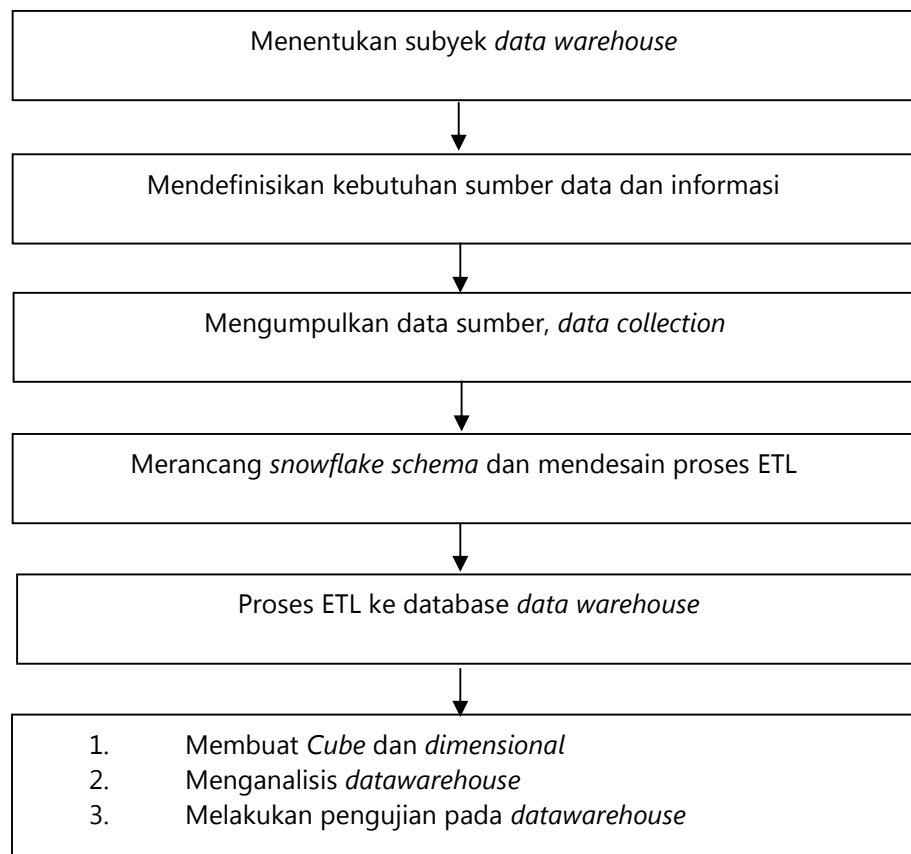
1.2. Rumusan Masalah

Berdasarkan paparan latar belakang di atas, maka permasalahan yang dapat dirumuskan adalah :

1. Bagaimana menyusun model data dengan menggunakan *snowflake schema* untuk membangun *data warehouse* Perpustakaan Perguruan Tinggi XYZ?
2. Bagaimana melakukan proses *Extraction, Transformation and Loading* (ETL) data dari basis data transaksional kedalam *data warehouse*?
3. Bagaimana menyusun serta menganalisisnya menggunakan *reporting system* untuk manajemen Perpustakaan?

1.3. Metodologi Penelitian

Proses penelitian ini dapat dilihat pada Gambar 1 di bawah ini:



2. Landasan Teori

2.1. Pengertian *Data warehouse*

Definisi *Data Warehouse* menurut William H Inmon (2005) adalah koleksi data yang berorientasi subyek (*subject-oriented*), terintegrasi (*integrated*), tidak dapat diubah (*nonupdateable*), memiliki periode waktu tertentu (*time variant*), untuk informasi yang bersifat historis dan dapat mendukung pengambilan keputusan. Sedangkan menurut Sean Kelly (1996), *Data Warehouse* adalah kumpulan data yang memiliki beberapa karakteristik, yaitu *separate, available, integrated, time-stamped, subject-oriented, non-volatile, dan accessible*.

Menurut Vincent Rainardi (2008), *Data warehouse* adalah sistem yang mengambil dan mengkonsolidasikan data secara periodik dari sistem sumber ke dalam table dimensi atau data penyimpanan yang sudah dinormalisasi. Data yang terkandung di dalamnya biasanya diperbarui dalam kurun waktu tertentu, tidak setiap kali transaksi terjadi dalam sistem sumber.

Data warehouse dibangun untuk mengatasi masalah teknis dan bisnis yang berkaitan dengan penggunaan data dan informasi untuk mengambil keputusan (Han dan Kamber, 2006). *Data warehouse* bukan suatu produk tetapi suatu lingkungan dimana user dapat menemukan informasi strategik. *Data warehouse* adalah kumpulan data-data logik yang terpisah dengan database operasional dan merupakan suatu ringkasan (Ponniah, 2001, h.14).

2.2. *Snowflake Schema*

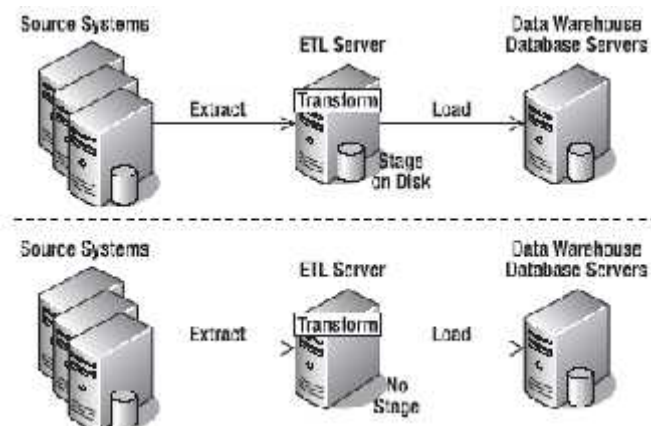
Snowflake schema adalah salah satu metode pemodelan data untuk *data warehouse* yang merupakan pengembangan dari *star schema*, yaitu metode normalisasi untuk tabel dimensi yang ada dalam *star schema* (Ponniah, 2001). Menurut Inmon (2005) *snowflake schema* adalah gabungan dari beberapa pemodelan data *star schema* yang dinormalisasi.

Dalam *snowflake schema*, setiap tabel dimensi dapat memiliki sub-tabel dimensi lagi. Hal ini bertujuan untuk meminimalkan data yang berlebihan (*redudancy data*) (Rainardi, 2008). 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. Contoh diagram *snowflake schema* dapat dilihat pada Gambar 2 di bawah ini:

2.3. ETL (*Extraction, Transformation, and Loading*)

ETL singkatan dari *Extraction, Transformation, and Loading*, adalah proses mengambil dan mengubah data dari sistem sumber dan memasukkannya ke dalam *data warehouse* (Rainardi, 2008. h.173). Dalam proses pengambilan data, data harus bersih agar didapat kualitas data yang baik. Contohnya ada nomor telepon yang invalid, ada kode buku yang tidak eksis lagi, ada beberapa data yang *null*, dan lain sebagainya. Pendekatan tradisional pada proses ETL mengambil data dari data sumber, meletakkan pada *staging area*, dan kemudian mentransform dan meng-load ke data warehouse.

Pada Gambar 2 di bawah ini dijelaskan alternative dalam proses ETL. Proses pertama menggunakan *staging area* dan proses kedua tidak menggunakan *staging area* dalam arti *staging area* tidak disimpan di dalam disk (Rainardi, 2008.h.174).



Gambar 2. Pendekatan dalam proses ETL

2.4. Perpustakaan Perguruan Tinggi (PPT)

Undang-undang No 2 Tahun 1989 tentang Sistem Pendidikan Nasional pada pasal 55 menyebutkan bahwa salah satu syarat untuk menyelenggarakan Perguruan Tinggi harus memiliki Perpustakaan. Dan dalam Undang-undang No. 43 Tahun 2007 tentang Perpustakaan dalam pasal 1, disebutkan bahwa Perpustakaan sebagai institusi pengelola koleksi karya tulis, karya cetak, dan/atau karya rekam secara profesional dengan sistem yang baku guna memenuhi kebutuhan pendidikan, penelitian, pelestarian, informasi, dan rekreasi para pemustaka. Sedangkan perpustakaan perguruan tinggi (PPT) merupakan unit pelaksana teknis (UPT) yang bersama-sama dengan unit lain melaksanakan Tri Dharma PT (Perguruan Tinggi) melalui menghimpun, memilih, mengolah, merawat serta melayani sumber informasi kepada lembaga induk khususnya dan masyarakat akademis pada umumnya. (Pedoman PPT, Jakarta: Dirjen DIKTI,

1994, hal. 3). Adapun yang termasuk dalam PT meliputi universitas, institut, sekolah tinggi, akademi, politeknik dan atau PT lain yang sederajat.

Perpustakaan perguruan tinggi adalah sebuah perpustakaan atau sistem perpustakaan yang dibangun, diadministrasikan dan didanai oleh sebuah universitas untuk memenuhi kebutuhan informasi, penelitian dan kurikulum dari mahasiswa, fakultas dan stafnya. Definisi ini menyatakan bahwa perpustakaan perguruan tinggi boleh berupa sebuah perpustakaan atau sistem perpustakaan. Sistem perpustakaan maksudnya adalah kemungkinan adanya berbagai perpustakaan baik perpustakaan pada tingkat universitas, fakultas, departemen/jurusan dan juga pada lembaga lain di bawah suatu perguruan tinggi. Artinya, pada sebuah perguruan tinggi dimungkinkan terdapat berbagai perpustakaan yang dimaksudkan untuk mendukung perguruan tinggi itu sendiri.

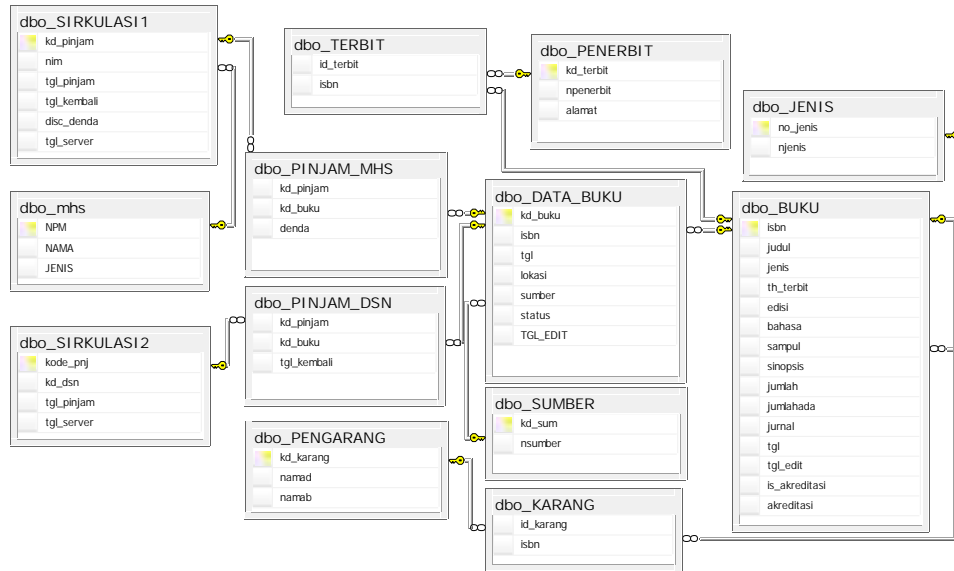
3. HASIL DAN PEMBAHASAN

3.1. Menentukan Subyek Data Warehouse

Dalam tahapan perencanaan proyek ini subyek utama dalam database perpustakaan adalah proses sirkulasi untuk peminjaman mahasiswa. Yaitu proses yang menjelaskan tentang peminjaman dan pengembalian koleksi perpustakaan khusus mahasiswa.

3.2. Mendefinisak Sumber Data dan Kebutuhan Informasi

Tahapan ini mendefinisikan semua sumber data yang akan digunakan dalam *data warehouse* dan mendefinisikan kebutuhan informasi yang harus dihasilkan oleh *data warehouse* melalui analisis *reporting*. Sumber data yang akan diolah untuk membangun *datawarehouse* diperoleh dari database transaksi perpustakaan berdasarkan periode tertentu. Data yang diambil adalah data transaksi perpustakaan sejak tahun 2003 sampai Januari 2013.



Gambar 3. Tabel Relasi Database Perpustakaan

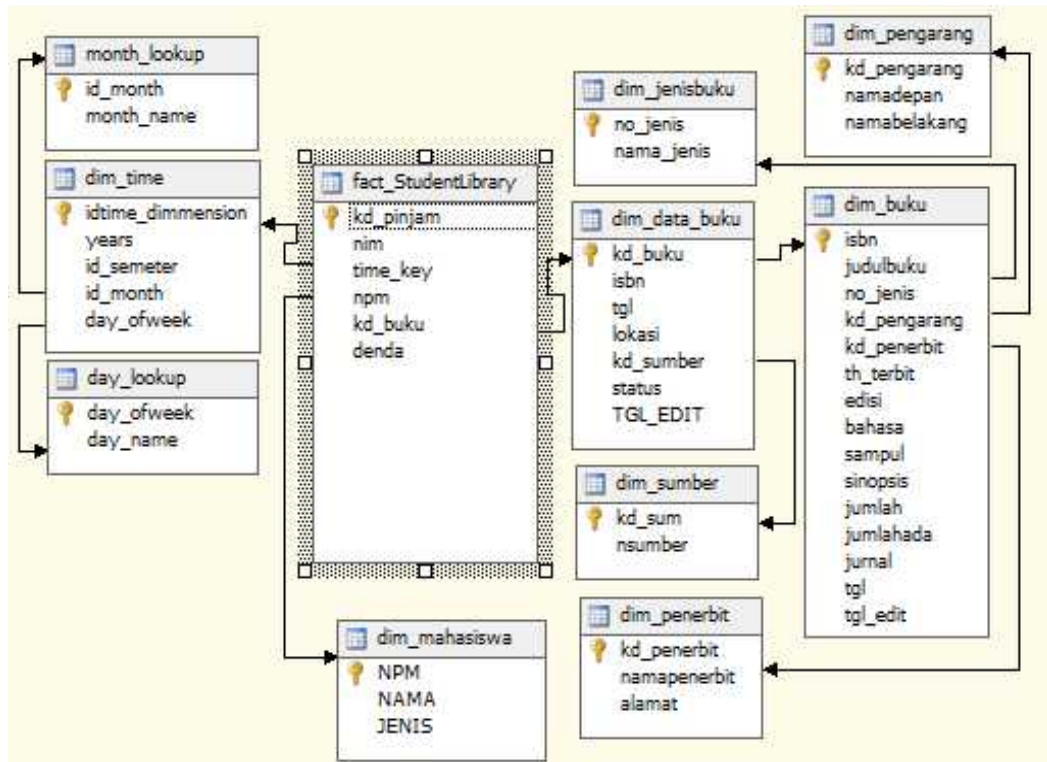
Berdasarkan hasil pengamatan terhadap perpustakaan beberapa kebutuhan informasi yang penting adalah sebagai berikut.

1. Analisis data terhadap buku-buku yang dipinjam oleh mahasiswa.
2. Data mahasiswa yang paling sering meminjam buku berdasarkan tabel multidimensi.
3. Informasi tentang jenis buku yang paling sering dipinjam.
4. Analisis *time series* atau *trend* proses sirkulasi.
5. Informasi tentang nama pengarang yang bukunya paling sering dipinjam.
6. Informasi tentang nama penerbit yang bukunya paling sering dipinjam.
7. Analisis sirkulasi yang berkaitan dengan tabel multidimensi.

3.3. Merancang Snowflake Schema

Dalam tahapan ini kegiatan yang dilakukan adalah membuat model data dimensional yang berbentuk *snowflake schema* dan dilanjutkan membuat rancangan *proses ETL*. Yang menjadi subyek utama dalam warehouse perpustakaan adalah sirkulasi peminjaman buku untuk mahasiswa yang akan menjadi pusat skema. Dalam perancangan ini diimplementasikan dalam bentuk tabel fakta yang diberi nama **Fact_StudentLibrary**. Sedangkan data yang lain diimplentasikan menjadi tabel dimensi maupun sub-dimensi yang semuanya terkait dengan tabel fakta.

Model rancangan *snowflake schema* tersebut dapat dilihat pada Gambar 4 berikut ini:



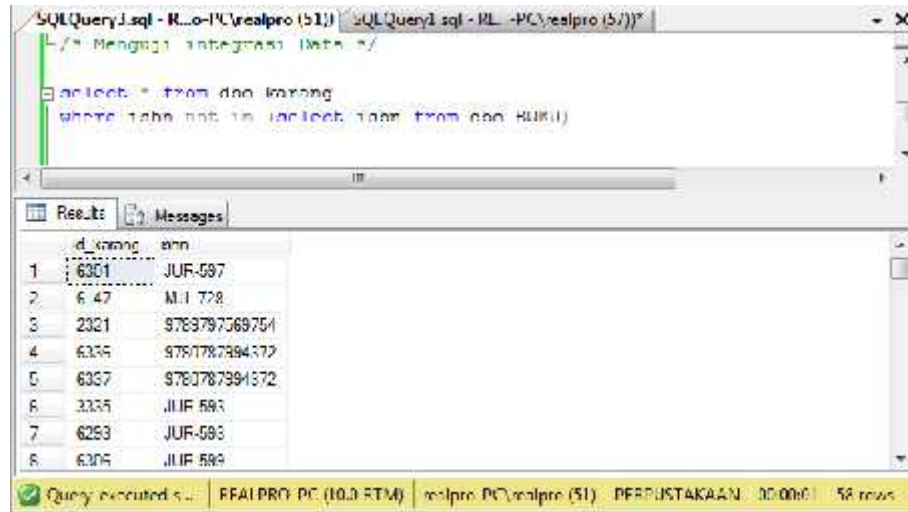
Gambar 4. Model Snowflake Schema Data Warehouse

3.4. Melakukan Proses ETL

Pada proses ETL dari database perpustakaan ke database *data warehouse* ada beberapa hal yang dilakukan. Secara garis besar terbagi menjadi 2 tahapan, yaitu:

3.4.1. Penghilangan Noise Data

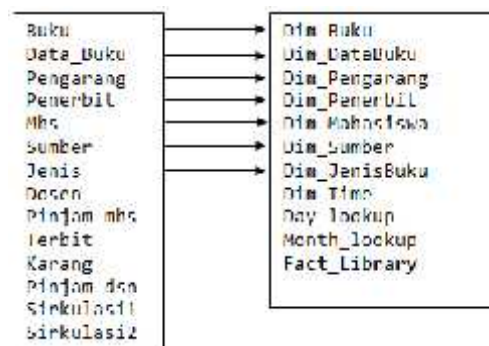
Sebelum proses ETL dilakukan perlu adanya proses menghilangkan *noise* yang ada pada database sumber atau database OLTP. Hal ini dilakukan untuk menjaga integritas data yang akan diolah. Setelah dilakukan proses analisis terhadap *database*, ternyata banyak sekali adanya *noise* ada pada database sumber. Sebagai contoh, terdapat 58 baris pada tabel *dbo_Karang* yang tidak konsisten. Ini bisa dilihat dari adanya data *isbn* pada tabel *dbo_Karang* padahal data *isbn* tersebut tidak ada pada tabel referensinya yaitu tabel *dbo_Buku*. Bukti proses penghapusan noise dapat dilihat pada Gambar 5 di bawah ini:



Gambar5. Pengujian Terhadap Konsistensi Data

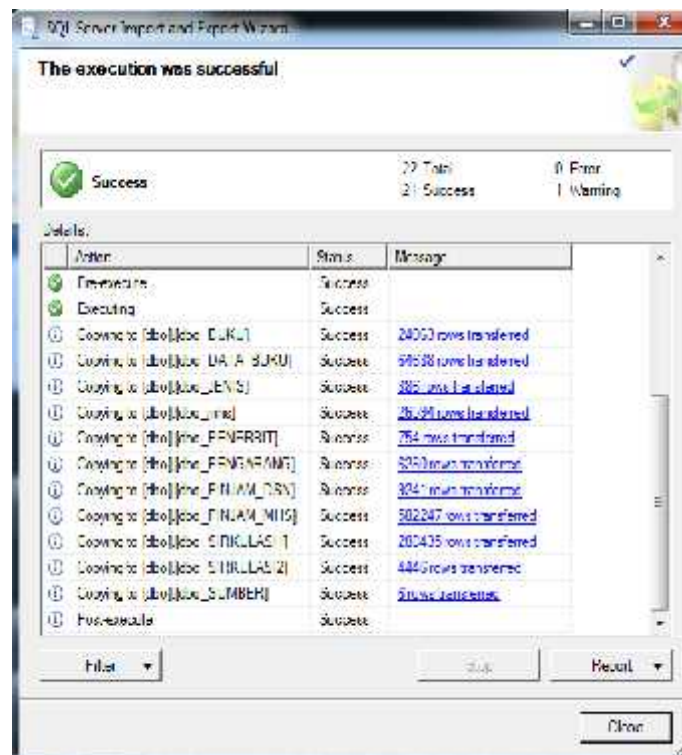
3.4.2. Proses Populasi Data

Proses ekstraksi data dari *database* operasional perpustakaan ke *database data warehouse* menggunakan suatu teknik populasi. Proses ini terdiri dari teknik untuk mempopulasikan data ke tabel dimensi dan tabel fakta. Rancangan proses populasi data dari tabel di *database* operasional tabel dimensi dan tabel fakta dapat dilihat pada Gambar 6 di bawah ini. Rancangan tersebut menggambarkan alir populasi tabel dimensi untuk tabel buku, tabel pengarang, tabel penerbit, tabel sumber, tabel jenis, tabel mahasiswa, dilakukan secara langsung dari *database* OLTP.



Gambar 6. Mapping Proses Populasi Data

Setelah proses mapping untuk semua tabel berhasil dilakukan, maka proses DTS akan berjalan sampai semua data terdistribusi sesuai dengan pemetaan tabel yang dilakukan. Indikasi keberhasilan proses DTS dapat dilihat pada Gambar 7 di bawah ini.



Gambar 7. Indikasi Proses DTS Berhasil

Proses selanjutnya adalah proses *loading* berkala pada setiap tabel dimension. Proses *loading* berkala dilakukan secara periodik berdasarkan waktu yang dibutuhkan sesuai dengan perkembangan data pada database OLTP. Proses ini dilakukan dengan menggunakan perintah *transact SQL* pada database *data warehouse*.

```

/* 1- Populasi Data dim_jenisbuku */
insert into dim_jenisbuku
select * from DBPerpus.dbo.dbo_JENIS
where DBPerpus.dbo.dbo_JENIS.no_jenis <> dim_jenisbuku.no_jenis

/* 2- Populasi Data dim_penerbit */
insert into dim_penerbit
select * from DBPerpus.dbo.dbo_PENERBIT
where DBPerpus.dbo.dbo_PENERBIT.kd_terbit <> dim_penerbit.kd_penerbit

/* 3- Populasi Data dim_Pengarang */
insert into dim_pengarang
select * from DBPerpus.dbo.dbo_PENGARANG
where DBPerpus.dbo.dbo_PENGARANG.kd_karang <> dim_pengarang.kd_pengarang

/* 4- Populasi Data dim_mahasiswa*/

```

```

insert into dim_mahasiswa
select * from DBPerpus.dbo.dbo_mhs
where DBPerpus.dbo.dbo_mhs.NPM <> dim_mahasiswa.NPM

/* 5- Populasi Data dim_Sumber */
insert into dim_sumber
select * from DBPerpus.dbo.dbo_SUMBER
where DBPerpus.dbo.dbo_SUMBER.kd_sum <> dim_sumber.kd_sum

/* 6- Populasi Data dim_buku*/
insert into dim_buku
select [isbn],[judul],[jenis]
,[th_terbit],[edisi]
,[bahasa],[sampul]
,[sinopsis],[jumlah]
,[jumlahada],[jurnal]
,[tgl],[tgl_edit]
FROM DBPerpus.dbo.dbo_BUKU
Where DBPerpus.dbo.dbo_BUKU.isbn <> dim_buku.isbn

/* 7- Populasi Data dim_data_buku*/
insert into dim_data_buku
select * from DBPerpus.dbo.dbo_DATA_BUKU
where DBPerpus.dbo.dbo_DATA_BUKU.kd_buku <> dim_data_buku.kd_buku

```

3.5. Analisis Hasil

Analisis pada *data warehouse* pada penelitian ini dilakukan dengan cara membuat suatu *reporting* sesuai dengan *business requirement*. Reporting dibangun dengan menjalankan perintah *transact SQL*.

Adapun beberapa *report* yang bisa dibuat adalah sebagai berikut:

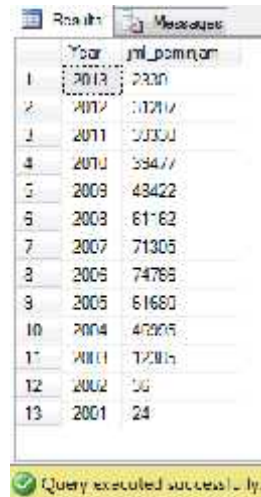
1. Laporan jumlah peminjaman pertahun.

Laporan ini memberikan informasi kepada manajemen Perpustakaan mengenai jumlah peminjaman pertahun.

```

--/* REPORTING PEMINJAMAN SETIAP TAHUN */
SELECT dim_time.Year,
COUNT(kd_pinjam) AS jml_peminjam
FROM fact_StudentLibrary JOIN dim_time
ON fact_StudentLibrary.time_key = dim_time.DateKey
GROUP BY dim_time.Year ORDER BY dim_time.Year DESC

```



	Year	jml_peminjam
1	2013	2330
2	2012	11211
3	2011	10300
4	2010	9947
5	2009	43422
6	2008	61162
7	2007	71305
8	2006	74766
9	2005	61660
10	2004	40005
11	2003	12005
12	2002	56
13	2001	24

Query executed successfully.

Gambar 8. Report Peminjaman setiap Tahun

2. Laporan buku terlaris.

Laporan ini memberikan informasi mengenai judul buku dan nama pengarangnya, dimana buku tersebut merupakan buku terlaris. Dengan melihat laporan ini pihak manajemen Perpustakaan bisa menentukan kebijakan pada saat proses pengadaan buku

```

/* REPORTING BUKU TERLARIS */
SELECT
dim_buku.isbn,
dim_buku.judul,
dim_pengarang.namadepan + dim_pengarang.namadepan AS Pengarang,
COUNT(dim_buku.isbn) AS jml_dipinjam
FROM dim_buku
INNER JOIN
fact_StudentLibrary ON dim_buku.isbn = fact_StudentLibrary.isbn
INNER JOIN
dim_pengarang ON fact_StudentLibrary.kd_karang = dim_pengarang.kd_karang
GROUP BY dim_buku.isbn, dim_buku.judul,
dim_pengarang.namadepan, dim_pengarang.namabelakang
ORDER BY jml_dipinjam DESC

```

3. Laporan hari paling banyak peminjam.

Laporan ini memberikan informasi mengenai hari teramai atau hari yang paling banyak proses peminjaman. Dari informasi ini pihak manajemen Perpustakaan bisa mengantisipasi saat-saat dimana hari yang paling banyak proses sirkulasinya sehingga mampu melayani dengan lebih baik. Misalnya penentuan jadwal waktu shift petugas perpustakaan harus berada

ditempat, jika ada kesibukan lain diupayakan tidak pada saat banyak proses sirkulasi, dan lain sebagainya.

```

/* REPORT HARI TERAMAI */
SELECT dim_date.DayofWeek AS hari,
COUNT(fact_StudentLibrary.kd_pinjam) AS jml_pinjam
FROM fact_StudentLibrary INNER JOIN
dim_time ON fact_StudentLibrary.time_key = dim_time.idtime_dimmension
GROUP BY dim_time.DayofWeek
ORDER BY jml_pinjam DESC

```

4. Laporan pengarang terlaris.

Laporan ini memberikan informasi mengenai nama pengarang terlaris, dimana buku-buku yang dikarangnya sering dipinjam oleh pengunjung perpustakaan.

```

/*REPORT PENGARANG TERLARIS */
SELECT dim_pengarang.namadepan + dim_pengarang.namabelakang AS pengarang,
COUNT(fact_StudentLibrary.kd_pinjam) AS jml_pinjam
FROM dim_buku
INNER JOIN
fact_StudentLibrary ON dim_buku.isbn = fact_StudentLibrary.isbn
INNER JOIN
dim_pengarang ON fact_StudentLibrary.kd_karang = dim_pengarang.kd_karang
GROUP BY dim_pengarang.namadepan, dim_pengarang.namabelakang
ORDER BY jml_pinjam DESC

```

5. Laporan penerbit terlaris.

Laporan ini memberikan informasi mengenai nama penerbit, dimana buku-buku yang diterbitkannya sering dipinjam oleh pengunjung perpustakaan.

```

--/* REPORT PENERBIT TERLARIS */
SELECT dim_penerbit.namapenerbit as [Nama Penerbit],
COUNT(fact_StudentLibrary.kd_pinjam) AS [Jumlah Peminjaman]
FROM fact_StudentLibrary INNER JOIN
dim_penerbit ON fact_StudentLibrary.kd_penerbit= dim_penerbit.kd_penerbit
GROUP BY dim_penerbit.namapenerbit
ORDER BY [Jumlah Peminjaman] DESC

```

	Nama Penerbit	Jumlah Peminjaman
1	Indonesian	1.254.011
2	Wahana Media Komunikasi	1.015.673
3	Indonesian	2.100.000
4	Radio - Indonesia Yogyakarta	2.050.000
5	Indonesian	1.015.000
6	Indonesian	1.015.000
7	Indonesian	8.025.000
8	Indonesian	5.000.000
9	Selamat Indonesia	8.750.000
10	Indonesian	8.445.000
11	Indonesian	8.145.000
12	Indonesian	1.015.000
13	PT. Telekomunikasi Indonesia	1.015.000
14	Wahana Komunikasi	5.000.000
15	Indonesian	5.000.000
16	Selamat Indonesia	5.017.000

Gambar 9. Report Peminjaman Setiap Penerbit

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Setelah melakukan identifikasi masalah, perancangan, implementasi, dan pengujian *prototype* sistem diperoleh beberapa kesimpulan sebagai berikut:

1. Proses membangun *datawarehouse* dengan *snowflake schema* melalui beberapa tahap yang berkesinambungan, yaitu merancang skema, melakukan proses ETL dan membuat *cube-data dimensional*.
2. Penyusunan *datawarehouse* pada tahap ETL (*Extract Transform Loading*) terkendala karena adanya *field foreign key* yang tidak konsisten. Banyak *field* kunci tamu yang *null*. Ini disebabkan karena pada awal penerapan sistem integritas data belum maksimal.
3. Untuk menjamin agar konsistensi *foreign key* terjaga, maka dilakukan penghapusan data *noise* tersebut. Karena integritas data dalam database yang dijadikan sumber untuk menyusun *datawarehouse* sangat penting. Pemodelan data dan pembangunan sistem basis data transaksional yang benar akan memudahkan dalam penyusunan *datawarehouse*. Begitu juga sebaliknya.

4.2. Saran

Berikut ini adalah kekurangan dalam penelitian ini yang diharapkan dapat menjadi bahan masukan atau saran untuk peneliti selanjutnya:

1. Penentuan *object* database sebagai data *source* harus diperhatikan secara matang. Database yang memiliki integritas data yang baik akan memudahkan dalam pembangunan *datawarehouse*.
2. Pengembangan *datawarehouse* dengan menggunakan Proses *Business Intelligence* akan sangat membantu untuk melakukan analisis lebih lanjut.
3. Dari analisis tersebut, selanjutnya bisa dilakukan penelitian untuk membangun *analysis service*.

5. DAFTAR PUSTAKA

- Amborowati, A. (2009). *Perancangan dan Pembuatan Data Warehouse pada Perpustakaan STMIK AMIKOM Yogyakarta*. Tesis. MTI Universitas Gadjah Mada. Yogyakarta.
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. Morgan Kaufmann. San Fransisco.
- Inmon, W. H. (2005). *Building the Data Warehouse - Fourth Edition*. J.Wiley. New York.
- Mandias, G. (2012). *Perancangan Data Warehouse untuk Analisis Data Mahasiswa Menggunakan Pemodelan Star Schema Di Universitas Klabat*. Tesis. MTI Universitas Gadjah Mada. Yogyakarta.
- Patel, A. R. (2012). Data Modeling Techniques For Data Warehouse. *ZENITH, International Journal of Multidisciplinary Research*, Vol.2 Issue 2, February 2012, ISSN 2231 5780 , 240-246.
- Ponniah, P. (2001). *Data Warehouse Fundamentals: a Comprehensive Guide for IT Professional*. J.Wiley. New York.
- Prasetyo, E. (2011). *Perancangan Data Warehouse Sistem Informasi Eksekutif (Studi Kasus Data Akademik Prodi Teknik Elektro FT UGM)*. Tesis. MTI Universitas Gadjah Mada. Yogyakarta.
- Rainardi, V. (2008). *Building a Data Warehouse with Examples in SQL Server*. Apress. New York.
- Yuadi, I. (2007). *Perpustakaan Digital:Paradigma, Konsep, dan Teknologi Informasi yang digunakan*. Diakses 12 September 2012. dari <http://journal.unair.ac.id/filerpdf/perpustakaan%20digital.pdf>