Desain Dan Implementasi Data *Warehouse* Pada Perpustakaan Daerah Provinsi Sumatera Selatan

Nia Adila & Andri

Fakultas Teknik Ilmu Komputer, Universitas Bina Darma

Email: niaadila04@gmail.com

Abstract

This research will focus on processing visitor data, member data, book borrowing data and book return data at the Regional Library of South Sumatra Province. The data is stored in the form of an excel file with a large amount of data, causing problems in the storage system such as data accumulation, data loss, no data analysis and delays in the reporting process. The process of data storage and data retrieval will be well integrated by building a data Warehouse at the Regional Library of South Sumatra Province. Data Warehouse is a system that contains several years of history and facilitates decision making. At the data Warehouse design stage using the Nine-Step method (Kimball, 2002), in this method there are nine steps in designing a data Warehouse, namely Process Selection, Grain Selection, Identification of dimensional adjustments, Fact Selection, Storage of initial calculations in the fact table, Reviewing the dimension table, selecting the database duration, tracking dimension changes, and prioritizing, querying the model and selecting the physical design. And the design and data processing process will use the Pentaho kettle and public Tableau applications, with the design and implementation of the data Warehouse, it is expected to help facilitate the reporting and analysis process for the Regional Library of South Sumatra Province.

Keywords: Library, Data Warehouse, Pentaho kettle, Tableau

1. Pendahuluan

Kemajuan teknologi dan informasi pada era globalisasi saat ini berkembang dengan pesat, hal tersebut berdampak pada tersedianya media *online* untuk memudahkan dalam mengakses informasi dan ilmu pengetahuan (Almas 2017). Namun tidak semua informasi yang ada pada internet (*online*) memiliki sumber yang jelas sehingga disinilah peran dari media informasi *offline* yaitu buku, buku memiliki sumber yang jelas dan masih diminati oleh para pembaca, dimana tempat yang menyediakan dan memiliki berbagai jenis buku yang lengkap dan bisa diakses oleh siapa pun adalah perpustakaan.

Perpustakaan merupakan salah satu sumber informasi yang dapat dimanfaatkan oleh masyarakat pengguna untuk memenuhi kebutuhan informasi (Hermawan and Hamzah 2017). Sehingga untuk memenuhi kebutuhan informasi tersebut pada saat ini pemerintah Indonesia mengelola langsung perpustakaan di disetiap daerah Provinsi, khususnya pada daerah Provinsi Sumatera Selatan yang memiliki tingkat pengunjung yang cukup tinggi. Dan memiliki kumpulan data pengunjung, data anggota, data peminjaman buku dan pengembalian buku yang banyak sehingga harus selalu di *entry* pada *Microsoft excel* terlebih dahulu dan di laporkan setiap bulan menggunakan sistem penyimpanan data, dengan banyaknya data yang ada Perpustakaan Daerah Provinsi Sumatera Selatan sering mengalami penumpukan data dan memiliki kendala pada proses pelaporan baik untuk analisa data untuk pengambilan keputusan.

Dengan adanya implementasi system penyimpanan data yang lebih besar dari *Microsoft excel* yang hanya memiliki jumlah *record* lebih dari 1.048.576 baris atau 16.384 kolom (ANGGAKARA 2019), selain *Microsoft excel* ada juga *mysql* dengan jumlah maksimal

untuk *type int* adalah 4.294.967.295 sedangkan untuk *type bigint* terbesar adalah 18.446.744.073.709.551.615 (Tipamedia 2018). Disinilah peran *database* dengan memiliki kapasitas penyimpanan yang lebih besar, Data diakumulasi dengan menambahkan *record* baru, bukannya dijaga tetap paling mutakhir dengan memperbarui *record* yang sudah ada dengan informasi baru, dioptimalkan untuk melakukan *Online Analytical Processing* (OLAP) dengan efisiensi *queri* lebih dari *petabyte* data. Sehingga akan terintegrasi lebih baik pada perpustakaan memberikan dampak yang baik untuk prose analisa data dan pelaporan kepada pimpinan perpustakaan untuk pencarian data pada perpustakaan maka sistem penyimpanan perpustakaan akan lebih mudah dan akurat dengan merancang dan membuat data *Warehouse*.

Teknologi data *Warehouse* dibutuhkan oleh sebagian besar organisasi atau perusahaan dan Instansi, yang mana memungkinkan untuk melakukan integrasi beberapa macam data lama dan data baru agar tidak terjadi duplikasi data, menganalisis berbagai data *history* yang berasal dari berbagai sistem dan aplikasi dalam satu tempat penyimpanan yang besar, data yang telah diintegrasikan dapat diolah dalam berbagai bentuk laporan sesuai dengan kebutuhan (Abdurahman 2017).

Konsep data *Warehouse* menjadi alternatif yang tepat, dimana data *Warehouse* digunakan untuk mengintegrasikan informasi data *history* lama dan data baru perpustakaan dengan tidak lagi mencari ringkasan informasi yang penting untuk membuat keputusan bisnis tanpa harus menjelajahi keseluruhan data. Oleh karena itu solusi dari permasalahan diatas yaitu dengan merancang Desain dan Implementasi Data *Warehouse* Pada Perpustakaan Daerah Provinsi Sumatera Selatan.

2. Tinjauan Review

2.1 Perpustakaan

Perpustakaan adalah suatu unit kerja dari suatu badan atau lembaga tertentu yang mengelola bahan-bahan pustaka, baik berupa buku-buku maupun bukan berupa buku (*nonbook material*) yang diatur secara sistematis menurut aturan tertentu sehingga dapat digunakan sebagai sumber informasi oleh setiap pemakainya (Silvana, Akbar, and Tifani 2017).

2.2 Implementasi

Implementasi adalah pelaksanaan atau penerapan perluasan aktivitas dan juga bisa diartikan sebagai aktivitas yang saling menyesuaikan (Haryati, Sudarsono, and Suryana 2015).

2.3 Data

Data merupakan sebuah fakta atau observasi mentah yang biasanya mengenai transaksi bisnis, dan ukuran data objektif dari atribut (karakteristik) dari entitas, seperti orang, tempat, benda, atau kejadian.

2.4 Hirarki Data

Secara tradisional data di organisasikan ke dalam suatu hirarki yang terdiri atas (Edi and Betshani 2012) :

1. Elemen data

Adalah satuan data terkecil yang tidak dapat dipecah lagi menjadi unit lain yang bermakna. Istilah lain untuk elemen data adalah *field*, kolom, *item*, dan atribut.

2. Record

Adalah gabungan sejumlah elemen data yang saling terkait. Dalam sistem basis data relasional, *record* biasa disebut dengan istilah tupel atau baris.

2. File

Himpunan seluruh *record* yang bertipe sama membentuk sebuah *file*. *File* dapat dikatakan sebagai kumpulan data *record* yang berkaitan dengan suatu subjek. Dalam sistem basis data relasional, *file* mewakili komponen yang disebut tabel atau *relationship*.

2.5 Database

Database atau Basis data adalah kumpulan informasi yang disimpan di dalam komputer secara sistematis sehingga dapat diperiksa menggunakan suatu program komputer untuk memperoleh informasi dari basis data tersebut. *Database* digunakan untuk menyimpan informasi atau data yang terintegrasi dengan baik di dalam komputer (Solichin 2010).

2.6 Data Mining

Data *mining* merupakan suatu proses penggalian data dan informasi yang besar yang belum diketahui sebelumnya namun dapat dipahami dan berguna dari *database* yang besar serta digunakan untuk membuat suatu keputusan bisnis yang penting.

2.7 Data Warehouse

Data *Warehouse* adalah koleksi data yang mempunyai sifat berorientasi subjek, terintegrasi, *time-variant*, dan bersifat tetap dari kumpulan data dalam mendukung proses pengambilan keputusan manajemen (Abdurahman 2017). Data *Warehouse* bersumber dari berbagai format, *software*, *platform* dan jaringan yang berbeda, data yang ada dalam data *warehouse* merupakan hasil dari proses transaksi perusahaan/organisasi sehari-hari. Data *warehouse* merupakan salah satu sistem pendukung keputusan yaitu dengan menyimpan data dari berbagai sumber, diorganisasikan dan dianalisa oleh para pengambil kebijakan (Andri and Tujni 2017). Data *Warehouse* mempunyai beberapa karakteristik, yaitu :

1) Subject Oriented (Berorientasi Subyek)

Data *Warehouse* dirancang untuk membantu kita dalam menganalisis data dalam jumlah yang banyak secara keseluruhan. Data yang disusun menurut subyek berisi hanya informasi yang penting bagi pemrosesan Decision Support Sistem (DSS). Informasi yang tersimpan dalam database di kelompokkan berdasarkan subyek tertentu, contohnya dalam kasus perpustakaan adalah anggota dan buku. Data-data di setiap subyek dirangkum ke dalam dimensi, misalnya periode waktu, sehingga dapat memberikan data historis untuk bahan analisis.

2) Terintegrasi

Data yang terdapat dalam data *Warehouse* dapat berasal dari beberapa sumber terpisah yang nantinya semua data tersebut akan disimpan ke dalam satu bagian yang sama dengan suatu format khusus dan konsisten. Sumber data *Warehouse* berasal dari *database* operasional (*internal source*) dan dari luar sistem (*external source*). Data *Warehouse* dapat menyimpan data-data yang berasal dari sumber - sumber yang

terpisah kedalam suatu format yang konsisten dan saling terintegrasi satu dengan lainnya.

3) *Time-variant* (Rentang Waktu)

Data di dalam data *Warehouse* mengandung data historis yang akan digunakan untuk analisis dan pengambilan keputusan. Data di dalam data *Warehouse* mempunyai karakter khusus berupa *time-series* dalam bentuk data historikal yang bervariasi terhadap waktu. Fungsi ini dapat digunakan untuk melakukan trend analysis dari data tersebut.

4) *Non-volatile* (Bersifat tetap)

Data-data yang sudah terdapat di dalam data *Warehouse* tidak dapat dilakukan perubahan lagi. Data dalam data *Warehouse* tidak diupdate secara *real-time*/terus menerus (dengan menerapkan fungsi *insert*, *update*, *delete*) seperti dalam sebuah sistem OLTP, tetapi hanya dapat ditambah dan dilihat. Pada data *Warehouse* hanya ada dua kegiatan memanipulasi data yaitu : *loading* data (mengambil data), dan akses data (mengakses data *Warehouse*). Data di dalam data *Warehouse* secara periodik di unggah dalam jangka waktu yang sama (Seta, Wati, and Isnainiyah 2017).

2.8 Struktur Data Warehouse

(a) *Current Detail Data*

Merupakan level terendah dalam struktur dalam data *Warehouse*, menggambarkan data detail yang aktif pada saat ini dan keadaan yang sedang berjalan.

(b) *Older Detail Data*

Merupakan data *back-up* (cadangan) yang jarang diakses.

(c) *Lightly Summarized Data*

Merupakan data ringkasan dari *current detail* data. Pada tahap ini data belum dapat digunakan untuk pengambilan keputusan karena data belum bersifat total *summary*, yang artinya data masih bersifat detail.

(d) *Highly Summarized Data*

Merupakan data yang bersifat total *summary*. Pada level ini, data sangat mudah diakses, terutama untuk melakukan analisis perbandingan data berdasarkan urutan waktu dan analisis yang menggunakan data multidimensi.

(e) Metadata

Merupakan data tentang data, metadata merupakan gambaran tentang struktur, isi, kunci, indeks dan data. Metadata dapat dikelompokkan ke dalam :

1. *Technical* metadata, yang berisi informasi mengenai data *Warehouse* yang digunakan oleh *administrator* dan perancang data *Warehouse* Ketika melakukan pengembangan data *Warehouse*.
2. *Business* metadata, yang berisi informasi yang memberikan *user* suatu perspektif yang mudah dimengerti dari informasi yang tersimpan dalam data *Warehouse*.
3. Data *Warehouse operational information*, seperti data *history* (*snapshots*, *versions*), *ownership*, menelusuri jejak audit, pengguna data.

2.9 Extract, Transform, Load (ETL)

Extract, Transform, Load (ETL) merupakan proses yang sangat penting dalam data *Warehouse*, dengan *ETL* inilah data dari *operational* dapat dimasukkan ke dalam data

Warehouse. ETL juga dapat digunakan untuk mengintegrasikan data dengan sistem yang sudah ada sebelumnya. Tujuan *ETL* adalah mengumpulkan, menyaring, mengolah, dan menggabungkan data-data yang relevan dari berbagai sumber untuk disimpan ke dalam data *Warehouse*. Hasil dari proses *ETL* adalah dihasilkannya data yang memenuhi kriteria data *warehouse* seperti data yang historis, terpadu, terangkum, statis, dan memiliki struktur yang dirancang untuk keperluan proses analisis.

a. *Extract*

Langkah pertama pada proses ETL adalah mengekstrak data dari sumber - sumber data. Kebanyakan proyek data *Warehouse* menggabungkan data dari sumber-sumber yang berbeda.

b. *Transform*

Tahapan transformasi menggunakan serangkaian aturan atau fungsi untuk mengekstrak data dari sumber dan selanjutnya akan dimasukkan ke data *warehouse*.

c. *Load*

Fase *load* merupakan tahapan yang berfungsi untuk memasukkan data ke dalam target akhir, yang biasanya ke dalam suatu data *warehouse*. Jangka waktu proses ini tergantung pada kebutuhan organisasi. Beberapa data *Warehouse* dapat setiap minggu mengisi keseluruhan informasi yang ada secara kumulatif, data diubah, Sementara data *warehouse* yang lain (atau bagian lain dari data *Warehouse* yang sama) dapat menambahkan data baru dalam suatu bentuk yang historikal, contohnya setiap jam (Methodology and Wijaya 2017).

2.10 Online Analytical Processing (OLAP)

Online Analytical Processing (OLAP) merupakan sebuah kategori *software* yang memungkinkan *analist*, manajer dan eksekutif untuk mendapat keuntungan dari dalam data secara cepat, konsisten, dan interaktif dengan berbagai kemungkinan yang ada pada pandangan terhadap informasi yang ditransformasi dari data mentah ke dalam bentuk nyata yang dapat dipahami oleh user. Umumnya *OLAP* meliputi aktivitas - aktivitas seperti pembangkitan *query* (*query builder*), permintaan laporan khusus dan grafik, dan melakukan analisis statistik. *OLAP* menyediakan kemampuan pemodelan dan visualisasi untuk kumpulan data besar yang diambil dari database operasional dan lebih seringnya diambil dari data *warehouse* sendiri (Kurniawan 2011).

2.11 Pentaho Kettle

Pentaho Kettle merupakan perangkat lunak (*software*) *open source* yang dikeluarkan oleh Pentaho. Aplikasi *Pentaho* dikembangkan oleh *Pentaho corp* yang berpusat di Orlando Amerika Serikat. Software *Pentaho Kettle* dapat digunakan sebagai *tools* untuk mengintegrasikan data. *Pentaho Kettle* menyediakan fasilitas untuk melakukan proses ETL (*Extraction, Transformation dan Loading*)(Andri and Tujni 2017).

2.12 Multi-Dimensional Modeling

Multi dimensional Modeling adalah pendekatan dari data *warehouse* yang dibangun berdasarkan metode *On-Line Analytical Processing (OLAP)* yang memiliki konsep dimensi, hierarki, *level*, dan anggota. Konsep multidimensional *database* dapat melakukan analisa dari berbagai sudut dimensi, berbeda dengan konsep *database* relasional. Model ini biasa digunakan dalam data *warehouse* karena memiliki konsep

banyak dimensi atau perspektif pengukuran bisnis atau fakta-fakta (Nugroho, Tambotoh, and Hoetama 2008). Dalam multi dimensional modeling terdapat beberapa tahapan seperti berikut ini :

a) *Fact*

Fact adalah suatu koleksi dari relasi *data-data items*, terdiri dari ukuran-ukuran dan konteks *data*. Setiap *fact* biasanya merepresentasikan sebuah bisnis item, suatu transaksi bisnis, atau sebuah kejadian yang dapat digunakan dalam analisis bisnis atau proses bisnis. Dalam *data mart*, *fact* di implementasikan dalam tabel dasar dimana semua *data numerik* dan disimpan.

b) *Dimensions*

Dimensions adalah suatu koleksi dari anggota atau unit-unit *data* dengan tipe yang sama. Dalam sebuah diagram, suatu dimensi biasanya direpresentasikan dengan suatu *axis*. Dalam dimensional model, semua *data* menunjukan *fact* tabel yang diasosiasikan dengan satu dan hanya satu member dari setiap *multiple dimensions*. Jadi dimensi menunjukan latar belakang kontekstual dari *fact*. Banyak proses analisis yang digunakan untuk menghitung (*quatify*) dampak dari dimensi pada *fact*

c) *Measure*

Suatu *measure* (ukuran) adalah suatu besaran (angka numerik) atribut dari sebuah *fact*, yang menunjukan *performance* atau *behavior* (tingkah laku) dari bisnis secara relatif pada suatu dimensi. Angka atau nomor yang ditunjukan disebut dengan *variable*. Sebagai contoh ukuran dari penjualan dalam bentuk uang, besarnya penjualan, jumlah pengadaan, biaya pengadaan, banyaknya transaksi dan lainnya. Suatu ukuran dijelaskan dengan kombinasi dari member dari suatu dimensi dan diletakkan dalam *fact*.

2.13 Tableau Public

Tableau adalah perangkat lunak *business intelligence* yang mudah untuk digunakan, terutama dalam hal membuat visualisasi data, analisis data, dan pelaporan. Cara penggunaannya cukup mudah karena menggunakan sistem *drag and drop*. *Tableau* dapat menggabungkan data dari berbagai sumber data seperti *spreadsheet*, *database*, *cloud data*, dan *big data* ke dalam satu program untuk digunakan dalam suatu analisis yang dinamis.

3. Metode Penelitian

3.1 Metode Pengumpulan Data

Metodologi perancangan *data warehouse* yang digunakan dalam penelitian ini menggunakan metode *Nine-Step Methodolty* (Kimball, 2002). Objek penelitian ini adalah perpustakaan Daerah Provinsi Sumatera Selatan. Metode pengumpulan data yang digunakan adalah sebagai berikut:

a. Dokumentasi

Pada tahap ini peneliti melakukan dakumentasi data yang terkait dengan perpustakaan yang meliputi data buku, data pengunjung perpustakaan, data peminjaman buku, dan data pengembalian buku.

b. Wawancara

Pada tahap wawancara ini peneliti mengajukan beberapa pertanyaan tentang proses penyimpanan data pengunjung dan peminjam buku pada pihak perpustakaan Daerah Provinsi Sumatera Selatan.

c. Studi Pustaka

Pada tahap ini melakukan pencarian dan pemahaman teori dan informasi yang mendukung proses penelitian yang berasal dari jurnal, buku, dan informasi secara *online*.

3.1.1 *Pentaho Kettle*

Pentaho Kettle merupakan perangkat lunak (*software*) *open source* yang dikeluarkan oleh Pentaho. Aplikasi *Pentaho* dikembangkan oleh *Pentaho corp* yang berpusat di Orlando Amerika Serikat. Software *Pentaho Kettle* dapat digunakan sebagai *tools* untuk mengintegrasikan data. *Pentaho Kettle* menyediakan fasilitas untuk melakukan proses ETL (*Extraction, Transformation dan Loading*)(Andri and Tujni 2017).

Ectract, Transform, Load (ETL) merupakan proses yang sangat penting dalam data *Warehouse*, dengan *ETL* inilah data dari *operational* dapat dimasukkan ke dalam data *Warehouse*. *ETL* juga dapat digunakan untuk mengintegrasikan data dengan sistem yang sudah ada sebelumnya. Tujuan *ETL* adalah mengumpulkan, menyaring, mengolah, dan menggabungkan data-data yang relevan dari berbagai sumber untuk disimpan ke dalam data *Warehouse*. Hasil dari proses *ETL* adalah dihasilkannya data yang memenuhi kriteria data *warehouse* seperti data yang historis, terpadu, terangkum, statis, dan memiliki struktur yang dirancang untuk keperluan proses analisis.

3.1.2 *Data Warehouse*

Data Warehouse adalah koleksi data yang mempunyai sifat berorientasi subjek, terintegrasi, *time-variant*, dan bersifat tetap dari kumpulan data dalam mendukung proses pengambilan keputusan manajemen (Abdurahman 2017). *Data Warehouse* bersumber dari berbagai format, *software*, *platform* dan jaringan yang berbeda, data yang ada dalam data *warehouse* merupakan hasil dari proses transaksi perusahaan/organisasi sehari-hari.

Data Warehouse mempunyai beberapa karakteristik, yaitu :

1. *Subject Oriented* (Berorientasi Subyek)

Data Warehouse dirancang untuk membantu kita dalam menganalisis data dalam jumlah yang banyak secara keseluruhan. Data yang disusun menurut subyek berisi hanya informasi yang penting bagi pemrosesan Decision Support Sistem (DSS). Informasi yang tersimpan dalam database di kelompokkan berdasarkan subyek tertentu, contohnya dalam kasus perpustakaan adalah anggota dan buku. Data-data di setiap subyek dirangkum ke dalam dimensi, misalnya periode waktu, sehingga dapat memberikan data historis untuk bahan analisis.

2. Terintegrasi

Data yang terdapat dalam data *Warehouse* dapat berasal dari beberapa sumber terpisah yang nantinya semua data tersebut akan disimpan ke dalam satu bagian yang sama dengan suatu format khusus dan konsisten. Sumber data *Warehouse* berasal dari *database* operasional (*internal source*) dan dari luar sistem (*external source*). *Data Warehouse* dapat menyimpan data-data yang berasal dari sumber - sumber yang

terpisah kedalam suatu format yang konsisten dan saling terintegrasi satu dengan lainnya.

3. *Time-variant* (Rentang Waktu)

Data di dalam data *Warehouse* mengandung data historis yang akan digunakan untuk analisis dan pengambilan keputusan. Data di dalam data *Warehouse* mempunyai karakter khusus berupa *time-series* dalam bentuk data historikal yang bervariasi terhadap waktu. Fungsi ini dapat digunakan untuk melakukan trend analysis dari data tersebut.

4. *Non-volatile* (Bersifat tetap)

Data-data yang sudah terdapat di dalam data *Warehouse* tidak dapat dilakukan perubahan lagi. Data dalam data *Warehouse* tidak diupdate secara *real-time*/terus menerus (dengan menerapkan fungsi *insert*, *update*, *delete*) seperti dalam sebuah sistem OLTP, tetapi hanya dapat ditambah dan dilihat. Pada data *Warehouse* hanya ada dua kegiatan memanipulasi data yaitu : *loading* data (mengambil data), dan akses data (mengakses data *Warehouse*). Data di dalam data *Warehouse* secara periodik di unggah dalam jangka waktu yang sama (Seta, Wati, and Isnainiyah 2017).

Data *warehouse* memiliki beberapa struktur sebagai berikut :

a) *Current Detail Data*

Merupakan level terendah dalam struktur dalam data *Warehouse*, menggambarkan data detail yang aktif pada saat ini dan keadaan yang sedang berjalan.

b) *Older Detail Data*

Merupakan data *back-up* (cadangan) yang jarang diakses.

c) *Lightly Summarized Data*

Merupakan data ringkasan dari *current detail* data. Pada tahap ini data belum dapat digunakan untuk pengambilan keputusan karena data belum bersifat total *summary*, yang artinya data masih bersifat detail.

d) *Highly Summarized Data*

Merupakan data yang bersifat total *summary*. Pada level ini, data sangat mudah diakses, terutama untuk melakukan analisis perbandingan data berdasarkan urutan waktu dan analisis yang menggunakan data multidimensi.

e) Metadata

Merupakan data tentang data, metadata merupakan gambaran tentang struktur, isi, kunci, indeks dan data. Metadata dapat dikelompokkan ke dalam :

- (1) *Technical* metadata, yang berisi informasi mengenai data *Warehouse* yang digunakan oleh *administrator* dan perancang data *Warehouse* Ketika melakukan pengembangan data *Warehouse*.
- (2) *Business* metadata, yang berisi informasi yang memberikan *user* suatu perspektif yang mudah dimengerti dari informasi yang tersimpan dalam data *Warehouse*.
- (3) Data *Warehouse operational information*, seperti data *history* (*snapshots*, *versions*), *ownership*, menelusuri jejak audit, pengguna data.

3.1.3 Tableau Public

Tableau adalah perangkat lunak *business intelligence* yang mudah untuk digunakan, terutama dalam hal membuat visualisasi data, analisis data, dan pelaporan. Cara

penggunaannya cukup mudah karena menggunakan sistem *drag and drop*. Tableau dapat menggabungkan data dari berbagai sumber data seperti *spreadsheet*, *database*, *cloud data*, dan *big data* ke dalam satu program untuk digunakan dalam suatu analisis yang dinamis.

3.2 Analisis Data

Sumber data yang digunakan pada penelitian ini adalah basis data Perpustakaan daerah Provinsi Sumatera Selatan. Informasi *database* yang ada berisi mengenai koleksi buku, data peminjaman, data pengembalian, dan data anggota perpustakaan pada bulan Januari – Maret 2021.

a. Metode Perancangan (Desain)

Metode perancangan data *warehouse* yang digunakan dalam penelitian ini menggunakan metode *Nine-Step* (Kimbal, 2002). Dalam metode ini terdapat sembilan langkah dalam merancang sebuah data *warehouse* yaitu *choosing the process*, *choosing the grain*, *identifying and conforming the dimensions*, *choosing the fact*, *storing pre-calculation in the fact table*, *rounding out the dimension tables*, *choosing the duration of database*, *tracking slowly changing dimension*, dan *deciding the query priorities and the query models*.

3.2.1 Desain Data Warehouse

Berikut desain dari data *warehouse* pada penelitian ini :

1. Pemilihan Proses (*Choose the process*).

Pada Perpustakaan Daerah Provinsi Sumatera Selatan memiliki proses bisnis yang meliputi: Penyerahan bebas Pustaka, pembelian buku, pendaftaran anggota, peminjaman buku, pengembalian buku.

2. Pemilihan Grain (*Choose the grain*)

Proses ini dilakukan setelah mengetahui proses bisnis yang ada pada perpustakaan. *Grain* ini akan dijadikan sebagai *fact table* dalam data *Warehouse*, berdasarkan proses bisnis yang ada berikut *grain* yang dihasilkan, jumlah buku, jumlah Mahasiswa, jumlah Pelajar, jumlah Umum dan jumlah buku yang dipinjam dan dikembalikan.

3. Identifikasi dari penyesuaian dimensi (*Identify and conform the dimensions*).

Langkah berikutnya adalah mengidentifikasi dimensi yang berhubungan dengan tabel fakta, dari identifikasi tersebut dimensi yang terlibat meliputi, a) Dimensi Anggota, merupakan anggota yang melakukan peminjaman buku pada perpustakaan. *Grain* yang terhubung dengan dimensi ini adalah jumlah anggota berdasarkan statusnya. b) Dimensi Buku, merupakan koleksi buku yang tersedia pada perpustakaan, dimensi buku memiliki hirarki kategori, pengarang dan penerbit buku. *Grain* yang berhubungan dengan dimensi ini yaitu jumlah buku yang di pinjam dan dikembalikan yang berdasarkan kategori. c) Dimensi Waktu, merupakan waktu terjadinya transaksi peminjaman buku dan pengembalian buku, *Grain* yang terlibat dengan dimensi ini yaitu jumlah peminjaman dan pengembalian buku.

4. Pemilihan Fakta (*Choose the fact*).

Langkah selanjutnya dalam perancangan data *Warehouse* adalah memilih *fact table* berdasarkan pemilihan *grain* sebelumnya. Berikut *fact table* peminjaman dan *fact table* pengembalian dapat dilihat pada gambar 1 dan gambar 2.

Gambar 1. *Fact Table* Peminjaman

Gambar 2. *Fact Table* Pengembalian

Dari Analisa didapat dua tabel fakta terdiri dari *fact table* peminjaman dan *fact table* pengembalian. Pada tabel fakta peminjaman dan tabel fakta pengembalian merupakan transaksi yang berisi atribut kunci yang berasal dari tabel-tabel dimensi yang telah dibuat sebelumnya. Model pada tabel menggunakan skema *snowflake*, skema ini memiliki percabangan pada tabel dimensinya dan dioptimalkan karena pertimbangan ruang penyimpanan, dimana keuntungan menggunakan skema *snowflake* memiliki kecepatan memindahkan data dari data OLTP kedalam metada, sebagai alat pengambil keputusan tingkat tinggi seluruh struktur dapat digunakan sepenuhnya.

5. Penyimpanan perhitungan awal pada tabel fakta (*Store pre-calculations in the fact table*). Pada tabel fakta peminjaman adalah total jumlah peminjaman buku berdasarkan waktu (hari, minggu, bulan dan tahun) dan agregasi fakta pengembalian buku adalah total pengembalian berdasarkan waktu (hari, minggu, bulan dan tahun).
6. Melihat kembali tabel dimensi (*Rounding out the dimension table*). Berikut tabel dimensi beserta atribut untuk membuat data *Warehouse* pada perpustakaan daerah provinsi Sumatera selatan :

Tabel 1. Dimensi Buku

Field	Type	Keterangan
IDBuku	Varchar (10)	Nomor Identitas Buku
IDKategori	Varchar (10)	Nomor Identitas Kategori
IDPengarang	Varchar (10)	Nomor Identitas Pengarang
IDPenerbit	Varchar (10)	Nomor Identitas Penerbit
Judul Buku	Varchar (100)	Judul Buku
Value	Varchar (30)	Nomor Value

Tabel 2. Dimensi Kategori

Field	Type	Keterangan
ID Kategori	Varchar (10)	Nomor Identitas Kategori
Nama Kategori	Varchar (40)	Kategori Buku Sosial dan Budaya, Ilmu Komputer)

Tabel 3. Dimensi Pengarang

Field	Type	Keterangan
ID Pengarang	Varchar (10)	NomorIdentitas Pengarang
Nama Pengarang	Varchar (40)	Nama Pengarang Buku

Tabel 4. Dimensi Penerbit

Field	Type	Keterangan
IDPenerbit	Varchar (10)	Nomor Identitas Penerbit
NamaPenerbit	Varchar (40)	Nama Penerbit Buku

Tabel 5. Dimensi Anggota

Field	Type	Keterangan
IDAnggota	Varchar (11)	Nomor Identitas Anggota
Nama Anggota	Varchar (100)	Nama Anggota
Status	Varchar (30)	Mahasiswa, Pelajar atau Umum

Tabel 6. Dimensi Waktu

Field	Type	Keterangan
IDWaktu	Varchar (10)	Nomor Identitas Waktu
Hari	Varchar (10)	Tanggal Peminjaman
Minggu	Varchar (10)	Mingguan
Bulan	Varchar (10)	Bulanan
Tahun	Varchar (10)	Tahunan

7. Pada Langkah berikutnya untuk pemilihan durasi basis data (*Decide the duration of the database and periodicity of updatation*). Dalam perancangan *data warehouse* perpustakaan ini durasi waktu yang digunakan selama 3 bulan dan data yang disimpan dalam *data warehouse* ini nantinya merupakan data 3 bulan terakhir yaitu data buku, data anggota, peminjaman dan pengembalian buku.
8. Menelusuri perubahan dimensi (*Track slowly the changing dimension*). Pendeskripsian yang tepat dari data pada tabel dimensi yang sudah lama harus dapat digunakan dengan histori transaksi yang lama. Mengamati perubahan dari dimensi pada tabel dimensi dapat dilakukan dengan tiga cara yaitu mengganti secara langsung pada tabel dimensi, membentuk *record* baru untuk setiap perubahan baru dan perubahan data yang membentuk kolom baru yang berbeda. Pada *data warehouse* ini telah dipilih cara kedua yaitu jika ada perubahan data, maka akan dibentuk *record* baru pada tabel dimensi. Data baru tersebut akan dimasukkan sebagai *record* baru tetapi *record* yang lama akan tetap ada, ini dilakukan agar semua proses yang terjadi dalam basis data dapat ditelusuri.
9. Penentuan prioritas, model *query* dan memilih *physical design* (*Decide the query priorities and the query models*). Pada tahap ini periode proses *Extract, Transform and Load* (ETL) dapat dilakukan sesuai dengan kebutuhan informasi oleh pihak pimpinan perpustakaan.

4. Hasil dan Pembahasan

4.1 Fact Table Peminjaman

Tools pentaho atau kettle digunakan untuk melakukan proses pembentukan tabel dimensi dan tabel fakta. kettle merupakan tools yang dapat digunakan untuk proses ETL (Extract Transform and Load). Pada gambar 3 menunjukkan proses pembentukan tabel fakta peminjaman buku. Pada proses pembentukan tabel fakta ini akan melibatkan beberapa tabel dimensi seperti, tabel dimensi buku, tabel dimensi waktu dan tabel dimensi anggota. tabel dimensi ini merupakan tabel detail yang menjelaskan informasi foreign key yang ada di tabel fakta.

Kolom-kolom yang ada dalam tabel dimensi ini dibuat secara hirarki sehingga memudahkan dalam proses queri. Sedangkan tabel fakta berisi berbagai nilai agregasi yang menjadi dasar pengukuran (measure) serta beberapa key yang terkait ke tabel dimensi yang akan menjadi sudut pandang dari measure tersebut. Measure merupakan sebuah kolom bernilai numerik yang dapat dilakukan proses agregasi seperti sum, max, avg dan lain-lain.

Gambar 3. Pembentukan Fact Table Peminjaman

Gambar 3 merupakan tahapan-tahapan dalam pembentukan tabel Fakta Peminjaman. Dalam lembar kerja pada gambar terdapat beberapa langkah yang digunakan untuk melakukan proses *Extract Transform dan Load* (ETL). Komponen-komponen ketel yang digunakan dalam pembentukan fakta tabel peminjaman terdiri dari input langkah, pencarian aliran langkah, langkah pilih nilai dan keluaran langkah.

Tahap pertama yang dilakukan adalah letakkan beberapa langkah dalam ketel yang digunakan untuk menghubungkan ke beberapa tabel dimensi yang terlibat dalam pembentukan tabel peminjaman, istilah ini disebut dengan ekstraksi data. Langkah berikutnya adalah melakukan proses integrasi terhadap tabel-tabel yang terlibat seperti tabel peminjaman, dimensi anggota, dimensi buku dan dimensi waktu.

Pada proses integrasi data juga terdapat proses transformasi yang digunakan untuk melakukan proses pengubahan ketika data dan nama kolom jika diperlukan. Tabel peminjaman merupakan tabel yang berisikan informasi mengenai peminjaman buku, tabel dimensi anggota merupakan tabel yang terbentuk dari tabel anggota yang berisikan informasi mengenai anggota perpustakaan, tabel dimensi buku adalah tabel yang berasal dari tabel buku yang berisikan informasi buku yang tersedia di perpustakaan, dan tabel dimensi waktu merupakan tabel yang berisikan data mengenai hierarki yang terdiri dari tahun, bulan dan tanggal.

Langkah terakhir adalah pembentukan tabel objek peminjaman yang dihasilkan dari penggabungan antara tabel peminjaman, tabel dimensi anggota, tabel dimensi buku dan dimensi waktu. Tabel fakta ini berisikan kumpulan yang menghubungkan ke tabel dimensi dan beberapa kolom yang dapat diukur (measure).

Gambar 4. Pembentukan Fact Table Pengembalian

Pada gambar 4 tahap pertama yang dilakukan adalah membuat beberapa step dalam *kettle* yang digunakan untuk menghubungkan ke beberapa tabel dimensi yang terlibat dalam pembentukan tabel fakta pengembalian istilah ini disebut dengan ekstraksi data.

Pada tahap ini akan dilakukan proses integrasi terhadap tabel-tabel yang terlibat seperti tabel pengembalian, dimensi anggota, dimensi buku dan dimensi waktu. Proses integrasi ini menggunakan step *stream lookup* yang ada dalam *pentaho*. Penggabungan antar tabel ini dapat dilakukan karena pada masing - masing tabel yang berelasi terdapat kolom kunci yang digunakan sebagai kunci penghubung antar tabel yang berelasi.

Langkah terakhir adalah pembentukan tabel fakta pengembalian_buku yang dihasilkan dari penggabungan antara tabel peminjaman, tabel dimensi anggota, tabel dimensi buku dan tabel dimensi waktu. Tabel fakta ini merupakan tabel transaksi yang berisikan kolom - kolom kunci yang berasal dari tabel - tabel dimensi yang berguna untuk mendapatkan informasi detail dari tabel-tabel dimensi. Proses pembentukan tabel fakta peminjaman dan fakta pengembalian_buku bertujuan untuk membentuk sebuah data *Warehouse* yang dapat digunakan untuk menyimpan data peminjaman buku dan data pengembalian buku perpustakaan.

4.2 Implementasi Tableau Public

Tableau adalah perangkat lunak *business intelligence* yang mudah untuk digunakan, terutama dalam hal membuat visualisasi data, analisis data, dan pelaporan. Cara penggunaannya cukup mudah karena menggunakan sistem *drag and drop*. *Tableau* dapat menggabungkan data dari berbagai sumber data seperti *spreadsheet*, *database*, *cloud data*, dan *big data* ke dalam satu program untuk digunakan dalam suatu analisis yang dinamis.

Penerapan *tableau public* akan didapatkan hasil berupa *dashboard* Peminjaman, pengembalian, jumlah buku dan jumlah anggota, dan akan melalui proses pembentukan *worksheet* menggunakan data dari data *warehouse* dw_perpustakaan seperti berikut ini.

Gambar 5. Tampilan *Dashboard*

Gambar 5 merupakan hasil dari pembuatan *dashboard* yaitu sebagai informasi jumlah peminjaman, pengembalian, jumlah buku, jumlah anggota dan jumlah anggota perstatus. Setelah membuat *dashboard* diatas, maka dilanjutkan dengan menganalisa hasil persentase setiap jumlah peminjaman, pengembalian, jumlah buku, jumlah anggota dan jumlah anggota perstatus pada bulan januari – maret 2021. Dan pada *dashboard* informasi

yang ada pimpinan perpustakaan bisa menggunakannya sebagai alat pengambil keputusan dengan membanding dan mengetahui jumlah data tersebut di setiap waktu, seperti berikut ini:

a. *Layout Parameter*

Gambar 6. Parameter Peminjaman dan Pengembalian

Parameter *dashboard* Perpustakaan Daerah Provinsi Sumatera Selatan pada gambar 6 memiliki 1(satu) selektor, yaitu selektor jumlah pinjaman dan pengembalian buku perminggu. Parameter dapat di ubah-ubah untuk dapat menampilkan grafik data yang berbeda berdasarkan selektor.

b. *Layout Line Chart* Persentase Peminjaman

Gambar 7. Grafik *Line Chart* Persentase Peminjaman

Dengan melihat grafik seperti pada gambar 7 user atau pimpinan dapat melihat informasi persentase Jumlah transaksi peminjaman. Grafik diatas dapat di ubah tampilan datanya berdasarkan *fact_peminjaman* dan *dimensi_waktu*. Sebagai contoh grafik diatas menunjukkan persentase minggu ke 3 dengan jumlah transaksi tertinggi yaitu 50 transaksi dan pada minggu ke 10 mengalami penurun yaitu hanya 13 transaksi pada bulan januari – maret tahun 2021.

c. *Layout Line Chart* Persentase Pengembalian

Gambar 8. Grafik *Line Chart* Persentase Pengembalian

Dengan melihat grafik seperti pada gambar 8 user atau pimpinan dapat melihat informasi persentase Jumlah transaksi peminjaman. Grafik diatas dapat di ubah tampilan datanya berdasarkan *fact_pengembalian* dan *dimensi_waktu*. Sebagai contoh grafik diatas menunjukkan persentase minggu ke 3 dengan jumlah transaksi tertinggi yaitu 24 transaksi dan pada minggu ke 10 mengalami penurun yaitu hanya 6 transaksi pada bulan januari – maret tahun 2021.

d. *Layout* Persentase Jumlah Buku

Gambar 9. Persentase Jumlah Buku

Dengan melihat grafik seperti pada gambar 9 user atau pimpinan dapat melihat informasi persentase jumlah keseluruhan transaksi buku yaitu 400 buku pada bulan januari – maret tahun 2021.

e. *Layout* Persentase Jumlah Anggota

Gambar 10. Persentase Jumlah Anggota

Dengan melihat grafik seperti pada gambar 4.33 user atau pimpinan dapat melihat informasi persentase jumlah keseluruhan anggota yaitu 437 pada bulan januari – maret tahun 2021.

f. *Layout* Persentase Jumlah Anggota Perstatus

Gambar 11. Persentase Jumlah Anggota Perstatus

Dengan melihat grafik seperti pada gambar 4.34 user atau pimpinan dapat melihat informasi persentase jumlah Anggota perstatus yaitu mahasiswa 369, pelajar 21 dan umum 46 pada bulan januari – maret tahun 2021.

5. Kesimpulan

Kesimpulan dari penelitian ini diambil dari pembahasan pada bab-bab yang sudah di bahas sebelumnya Desain dan Implementasi Data *Warehouse* Pada Perpustakaan Daerah Provinsi Sumatera Selatan adalah sebagai berikut :

1. Penelitian yang dilakukan menggunakan data transaksi peminjaman buku dan pengembalian buku.
2. Data *warehouse* dirancang menggunakan proses ETL menggunakan aplikasi *Pentaho*.
3. Penelitian yang dilaksanakan menghasilkan *dashboard business intelligence* transaksi peminjaman buku, pengembalian buku, jumlah buku dan jumlah anggota

- Perpustakaan Daerah Provinsi Sumatera Selatan yang menampilkan *dashboard* informasi berbentuk *chart* atau grafik dengan menggunakan aplikasi *Tableau public*.
4. *Dashboard Business Intelligence* yang dihasilkan terdiri dari *fact_peminjaman*, *fact_pengembalian*, dimensi waktu, dimensi buku dan dimensi anggota.

Referensi

- Abdurahman, Muhdar. 2017. "IJIS Indonesian Journal on Information System ISSN 2548-6438." *IJIS-Indonesia Journal on Information System* 2 (1): 18–26.
- Almas, Hanafi. 2017. "Manajemen Sistem Informasi Di Perpustakaan Smk Negeri 3 Malang." *BIBLIOTIKA: Jurnal Kajian Perpustakaan Dan Informasi* 1 (1): 91–100. <https://doi.org/10.17977/um008v1i12017p091>.
- Andri, A., and B. Tujni. 2017. "Perancangan Data Warehouse Perpustakaan." *Jurnal Ilmiah Matrik*, no. 3: 101–10. <https://doi.org/10.33557/jurnalmatrik.v19i2.378>.
- ANGGAKARA. 2019. "Jumlah Kolom Dan Baris Excel." 2019. <https://anggakarasahare.blogspot.com/2019/03/berapa-jumlah-maksimum-kolom-baris.html#:~:text=Jumlah maksimum baris %26 kolom di,576 baris dan 16.384 kolomdata>.
- Edi, Doro, and Stevalin Betshani. 2012. "Analisis Data Dengan Menggunakan ERD Dan Model Konseptual Data Warehouse." *Jurnal Informatika* 5 (1): 71–85.
- Haryati, Siska, Aji Sudarsono, and Eko Suryana. 2015. "Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu)." *Jurnal Media Infotama* 11 (2): 130–38.
- Hermawan, Herry, and Radja Erland Hamzah. 2017. "Jurnal Kajian." *Jurnal Kajian Perpustakaan Dan Informasi* 1 (2): 42–57.
- Kurniawan, Novianto Budi. 2011. "Desain Dan Implementasi Data Warehouse Studi Kasus Pemetaan Daerah Rawan Bencana Provinsi Sumatera Barat," 1–16.
- Methodology, Dengan Kimball Nine-step, and Ganda Wijaya. 2017. "Perancangan Data Warehouse Nilai Mahasiswa Dengan Kimball Nine-Step Methodology." *Jurnal Informatika (JI) UBSI* 4 (1): 1–11.
- Nugroho, Radityo Adi, Johan Tambotuh, and Tony Justinus Hoetama. 2008. "Aplikasi Data Warehouse Untuk Analisis Penjualan Mobil Berbasis Multidimensional Modeling (MDM) Dan Star Schema Design (Studi Kasus PT. Asco Automotive)." *Jurnal Teknologi Informasi-Aiti* 5 (2): 101–200.
- Program, Pada, and Ilmu Perpustakaan. 2019. "NIM: 1564400023 Diajukan Untuk Memenuhi Salah Satu Persyaratan Guna Memperoleh Gelar Sarjana Ilmu Perpustakaan (S.IP) Pada Program Ilmu Perpustakaan."
- Seta, Henki Bayu, Theresia Wati, and Ika Nurlaili Isnainiyah. 2017. "Data Warehouse Development for UPN 'Veteran' Jakarta Library (Perancangan Data Warehouse Pada Perpustakaan UPN 'Veteran' Jakarta)." *Journal Pekommas* 2 (2): 161. <https://doi.org/10.30818/jpkm.2017.2020206>.
- Silvana, Meza, Ricky Akbar, and Rahayu Tifani. 2017. "Penerapan Dashboard System Di

Perpustakaan Universitas Andalas Menggunakan Tableau Public.” *Seminar Nasional Sains Dan Teknologi 2017*, no. November: 1–6.

Solichin, Achmad. 2010. “MySQL Dari Pemula Hingga Mahir.” *Universitas Budi Luhur, Jakarta*, no. January 2010: 1–117.

Suni, Eugenius Kau. 2018. “Analisis Dan Perancangan Data Warehouse Untuk Mendukung Keputusan Redaksi Televisi Menggunakan Metode Nine-Step Kimball.” *Jurnal Teknik Informatika* 11 (2): 197–206. <https://doi.org/10.15408/jti.v11i2.8560>.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)