

PERANCANGAN DATA WAREHOUSE UNTUK MENDUKUNG PERENCANAAN PEMASARAN PERGURUAN TINGGI

Agung Prasetyo¹, Bambang Soedijono W², Armadyah Amborowati³

^{1,2,3}Magister Teknik Informatika

STMIK Amikom Yogyakarta

Email: pras@amikompurwokerto.ac.id¹, bambang.s@amikom.ac.id²,
armadyah.a@amikom.ac.id³

ABSTRAK

Salah satu indikasi perguruan tinggi yang besar adalah dilihat dari jumlah mahasiswa di perguruan tinggi tersebut. Karenanya, mahasiswa baru merupakan salah satu sumber daya yang menentukan berjalannya sebuah perguruan tinggi. Setiap tahunnya STMIK Amikom Purwokerto selalu melakukan penerimaan calon mahasiswa. Data mahasiswa baru tersebut sangat berguna bagi bagian pemasaran sebagai informasi untuk evaluasi kegiatan pemasaran berikutnya. Dengan dibangunnya data warehouse dan aplikasi OLAP dengan menggunakan aplikasi Pentaho Data Integration/Kettle sebagai perangkat ETL dan Pentaho Workbench yang merupakan Online Analytical Processing (OLAP) sebagai pengolah database, manajemen di STMIK Amikom Purwokerto bisa mengambil beberapa informasi misalnya; banyak jumlah pendaftar per-periode/gelombang, per/asal sekolahnya, per/asal sumber informasi yang diperoleh calon mahasiswa baru, serta tren minat terhadap jurusan yang dipilih oleh calon mahasiswa baru. Data warehouse mampu menganalisis data transaksi, mampu memberikan laporan yang dinamis dan mampu memberikan informasi dalam berbagai dimensi tentang penerimaan mahasiswa baru di STMIK Amikom Purwokerto.

Kata Kunci: Data Warehouse, OLAP, Pentaho, Penerimaan Calon Mahasiswa.

ABSTRACT

One indication of the large universities is seen from the number of students at the university. Therefore, new students is one resource that determines the passage of a college. Each year STMIK Amikom Purwokerto always do admissions. The new student data is very useful for the marketing department as information for the evaluation of subsequent marketing activities. With the construction of data warehouse and OLAP applications using application Pentaho Data Integration / ETL Kettle as devices and Pentaho Workbench is an Online Analytical Processing (OLAP) database as the processor, the management at STMIK Amikom Purwokerto could take some information, for example; lot number of applicants per-period / wave, per / her school of origin, per / origin obtained

resources for new students, as well as the trend of interest in the subject chosen by the prospective new students. Data warehouses are able to analyze transaction data, is able to provide a dynamic report and to provide information on the various dimensions of the new admissions in STMIK Amikom Purwokerto.

Keywords: Data Warehouse, OLAP, Pentaho, Admissions.

PENDAHULUAN

Salah satu indikasi perguruan tinggi yang besar adalah dilihat dari jumlah mahasiswa di perguruan tinggi tersebut. Karena mahasiswa baru merupakan salah satu sumber daya yang menentukan berjalannya sebuah perguruan tinggi. Setiap tahunnya STMIK AMIKOM Purwokerto selalu melakukan penerimaan calon mahasiswa. Data calon mahasiswa baru tersebut disimpan dalam suatu aplikasi SIA PMB dalam bentuk data SQL server. Selama ini, data calon mahasiswa baru tersebut tidak terkelola dengan baik. Padahal data calon mahasiswa baru tersebut sangat berguna bagi bagian pemasaran sebagai informasi untuk evaluasi kegiatan pemasaran berikutnya. Manajemen di STMIK AMIKOM Purwokerto bisa mengambil beberapa informasi misalnya; seberapa luas sebaran asal kota calon mahasiswa dilihat dari asal sekolahnya, seberapa banyak jumlah pendaftar per-periode/gelombang dilihat dari sumber asal informasi yang diperoleh calon mahasiswa baru tentang STMIK AMIKOM Purwokerto, tren minat terhadap jurusan yang dipilih oleh calon mahasiswa baru berdasarkan informasi dari mahasiswa.

Berdasarkan hal tersebut, maka di manajemen STMIK AMIKOM Purwokerto merasa perlu peningkatan penggunaan teknologi informasi untuk menganalisa data historikal mahasiswa baru. Salah satu penerapan penggunaan teknologi informasi tersebut menggunakan data warehouse yang terintegrasi dengan (On-Line Analytical Processing) OLAP yang dapat digunakan untuk menganalisis informasi yang berpotensi menunjang strategi pemasaran yang efektif dan efisien. Data warehouse memiliki kemampuan untuk melakukan analisa terhadap data transaksi yang telah lalu dan juga bisa memberikan laporan yang

bersifat dinamis dan bisa dilihat dari berbagai dimensi, sehingga dapat digunakan untuk mendukung proses analisis bagi para pengambil keputusan.

Dalam penelitian ini dibangun data warehouse dan aplikasi OLAP data penerimaan mahasiswa baru STMIK AMIKOM Purwokerto dengan menggunakan aplikasi Pentaho Data Integration/Kettle sebagai perangkat ETL dan Pentaho Workbench yang merupakan Online Analytical Processing (OLAP) sebagai pengolah database yang mampu menganalisis data transaksi dan mampu memberikan laporan yang dinamis dan mampu memberikan informasi dalam berbagai dimensi tentang penerimaan mahasiswa baru di STMIK AMIKOM Purwokerto.

Penggunaan aplikasi Pentaho Data Integration/Kettle sebagai perangkat ETL dan Pentaho Workbench yang merupakan Online Analytical Processing (OLAP) sebagai pengolah database dalam pembuatan laporan dapat mempermudah dalam menampilkan informasi untuk menjawab pertanyaan-pertanyaan yang berhubungan dengan penerimaan mahasiswa baru. Misalnya berupa: “Bagaimana laporan calon mahasiswa baru per-jurusan berdasarkan per-gelombang, per-jurusan dan per-informasi, per-jurusan dan per-asal sekolah, per-gelombang, per-penghasilan orang tua?”

Dengan terjawabnya pertanyaan-pertanyaan tersebut, dapat membantu manajemen institusi pendidikan STMIK AMIKOM Purwokerto khususnya bagian marketing dalam pengambilan keputusan yang tepat misalnya strategi pemasaran, metode ceramah di sekolah tertentu yang disesuaikan dengan data historikal penerimaan mahasiswa baru dari sekolah tersebut, dan lain-lain.

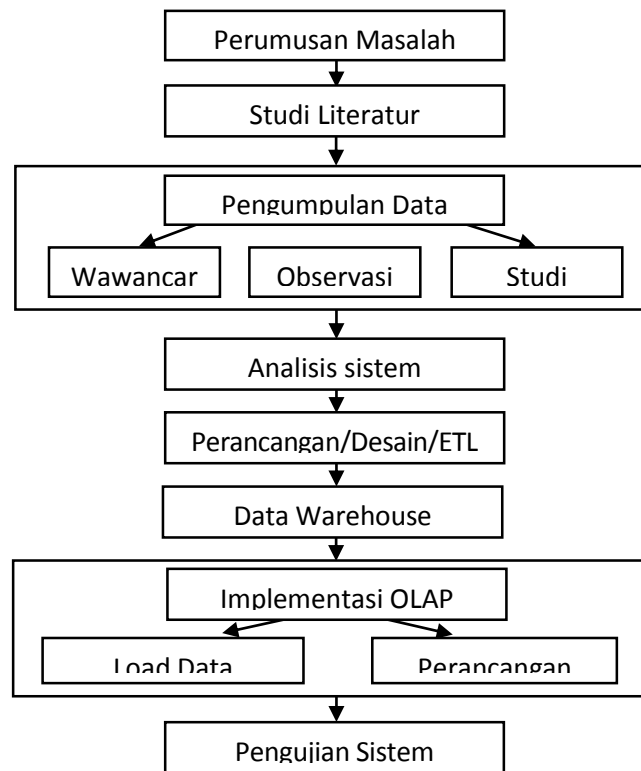
METODE PENELITIAN

A. Lokasi Penelitian

Penelitian dilakukan di STMIK Amikom Purwokerto. Data diperoleh dari bagian pemasaran STMIK Amikom Purwokerto

B. Alur Penelitian

Adapun tahapan proses penelitian dapat dilihat pada diagram alur proses Gambar 1 berikut ini:



Gambar 1. Alur Penelitian

C. Metode Penelitian

Metode penelitian yang digunakan termasuk dalam kategori riset terapan (Applied Research), dimana penelitian dilakukan dengan mengambil permasalahan yang terdapat di STMIK Amikom Purwokerto.

D. Metodologi Penelitian

Menurut Kimball ada sembilan tahap metodologi dalam perancangan database untuk data warehouse, yaitu:

a. Langkah 1: Pemilihan proses

Data mart yang pertama kali dibangun haruslah data mart yang dapat dikirim tepat waktu dan dapat menjawab semua pertanyaan bisnis yang penting

b. Langkah 2: Pemilihan sumber

Untuk memutuskan secara pasti apa yang diwakili atau direpresentasikan oleh sebuah tabel fakta.

c. Langkah 3: Mengidentifikasi dimensi

Set dimensi yang dibangun dengan baik, memberikan kemudahan untuk memahami dan menggunakan data mart, Dimensi ini penting untuk menggambarkan fakta-fakta yang terdapat pada tabel fakta.

d. Langkah 4: Pemilihan fakta

Sumber dari sebuah tabel fakta menentukan fakta mana yang bisa digunakan dalam data mart. Semua fakta harus diekspresikan pada tingkat yang telah ditentukan oleh sumber

e. Langkah 5: Menyimpan pre-kalkulasi di tabel fakta, Hal ini terjadi apabila fakta kehilangan statement

f. Langkah 6: Melengkapi tabel dimensi

Pada tahap ini ditambahkan keterangan yang lengkap pada tabel dimensi. Keteranganannya harus bersifat intuitif dan mudah dipahami oleh pengguna

g. Langkah 7: Pemilihan durasi database, Misalnya pada suatu perusahaan asuransi, mengharuskan data disimpan selama 10 tahun atau lebih

h. Langkah 8: Menelusuri perubahan dimensi yang perlahan. Ada tiga tipe perubahan dimensi yang perlahan, yaitu:

1. Tipe 1. Atribut dimensi yang telah berubah tertulis ulang
2. Tipe 2. Atribut dimensi yang telah berubah menimbulkan sebuah dimensi baru
3. Tipe 3. Atribut dimensi yang telah berubah menimbulkan alternatif sehingga nilai atribut lama dan baru dapat diakses secara bersama pada dimensi yang sama.

i. Langkah 9: Menentukan prioritas dan mode query, Pada tahap ini digunakan perancangan fisik.

E. Pengujian Sistem

Untuk mengetahui kinerja sistem dilakukan uji produk yang dilakukan oleh karyawan Bagian Marketing, Bagian Customer Service, panitia penerimaan mahasiswa baru, bagian IT, dan pimpinan. Atribut Uji Produk yang digunakan berdasarkan ISO 9126 yang mengidentifikasi enam karakteristik kualitas software utama yaitu, Fungsi (*functionality*), Keandalan (*reliability*), Kemudahan (*usability*), Efisien, (*efficiency*), *Maintainability*, *Portability*.

HASIL DAN PEMBAHASAN

A. Analisis Sistem

Dari analisis data, sumber data yang digunakan berasal dari sumber data SIA PMB yaitu tabel calonsiswa. Sumber data tersebut di konversi dalam Ms-excel.

Terdapat 2 (dua) sumber data dalam format Ms-excel, yaitu sumber data Pendaftaran dan Registrasi. Masing-masing tabel mewakili kegiatan transaksional yang terjadi dalam penerimaan calon mahasiswa baru.

permasalahan dan penyebab masalah yang muncul dalam Penerimaan Mahasiswa Baru (PMB) sebagai berikut:

Tabel 1. Hasil Identifikasi Permasalahan

No	Masalah yang dihadapi	Identifikasi penyebab masalah	Alternatif pemecahan masalah
1.	Data yang ada pada perguruan tinggi yang semakin besar dan banyak membutuhkan alokasi tempat penyimpanan	Belum tersedianya data warehouse.	Data warehouse dan aplikasi OLAP Penerimaan Mahasiswa Baru (PMB)
2.	Pengolahan data mahasiswa belum optimal khususnya pengolahan data yang berhubungan dengan analisis historikal mahasiswa baru	Belum tersedianya aplikasi OLAP yang digunakan dalam penerimaan Mahasiswa Baru (PMB)	
3.	Data penerimaan mahasiswa baru tidak dapat disajikan dalam bentuk multi dimensional		

Hasil analisa masalah dan identifikasi penyebab masalah yang ada dapat diambil suatu alternatif pemecahan masalah yaitu dengan membangun data

warehouse Penerimaan Mahasiswa Baru (PMB) yang dapat digunakan untuk menemukan pengetahuan maupun analisis data.

Alasan membangun data warehouse yaitu dengan tujuan untuk menyatukan data yang beragam ke dalam sebuah tempat penyimpanan dimana pengguna dapat dengan mudah menjalankan query (pencarian data), menghasilkan laporan, dan melakukan analisis sehingga dapat meningkatkan efektifitas pembuatan keputusan. Tanpa data warehouse untuk menyimpan informasi historikal, data disimpan di media statik atau di akumulasikan di database OLTP. Jika data diakumulasikan di database OLTP, maka akan membebani biaya query yang semakin lama semakin besar. Sehingga dapat mengganggu kinerja di database OLTP. Query dapat menjadi sangat kompleks terkait dengan skema database OLTP. Atas dasar tersebut diperlukan pengorganisasian dan konsolidasi data ke dalam sebuah data warehouse sehingga memungkinkan OLTP tetap dapat berjalan untuk transaksi yang tinggi. Selain itu query yang digunakan untuk analisis dapat lebih efisien dan sederhana.

B. Perancangan Sistem

Dalam Implementasi ini diperlukan rancangan data warehouse yang terdiri dari beberapa tahap agar perancangan data warehouse menjadi teratur. Tahap-tahap dalam perancangan data warehouse ini mengacu kepada Kimball yang dikutip oleh (Connolly, 2005), yaitu Nine-Step Methodology (kimball, 1996). Berikut penjelasan mengenai tahap-tahap perancangan dan pembuatan data warehouse:

a. Menentukan Proses (Choosing the process)

Proses pada penelitian ini adalah proses penerimaan mahasiswa baru. Pada penerimaan dimaksud adalah pada proses pendaftaran calon mahasiswa baru dan registrasi mahasiswa baru. Hasil data keluaran yang dicatat meliputi data diri calon mahasiswa, asal sekolah, prodi yang dipilih, waktu pendaftaran berdasarkan gelombang dan sumber informasi pendaftaran. Sumber data transaksi berupa data yang dikeluarkan oleh system berupa file Ms. excel.

b. Menentukan Grain (Choosing the grain)

Grain pada data transaksi penerimaan mahasiswa baru yang digunakan untuk merancang data warehouse ini adalah tabel fakta (fact_table).

c. Identifikasi dan Membuat Dimensi yang Sesuai (Identifying & Conforming the Dimensions)

Tahap berikutnya menjelaskan kegiatan pengidentifikasian hubungan dimensi dengan grain fakta. Menentukan Fakta (Choosing the facts), Setelah menentukan dimensi terkait, selanjutnya adalah memastikan tabel fakta yang diperoleh dalam proses grain.

d. Melengkapi Tabel Dimensi (Rounding out the dimension tables)

Tahapan ini menentukan kelengkapan tabel-tabel dimensi yang akan digunakan untuk membuat skema database. Kelengkapan tabel dimensi berisi data antara lain nama field, tipe data dan panjang data.

e. Memilih durasi dari database (Choose the durations of the database)

Pemilihan durasi data histori dapat dilakukan sesuai dengan kebutuhan informasi pihak eksekutif. Pada umumnya semakin banyak data yang disimpan dalam data warehouse, maka semakin lengkap pula informasi yang bisa dihasilkan. Walaupun sebenarnya semakin lama data yang disimpan juga menimbulkan beberapa permasalahan, antara lain permasalahan menginterpretasikan file lama dan versi dimensinya adalah versi lama bukan versi terbaru dan akhirnya hanya menjadi data sampah yang tidak dapat dimanfaatkan. Sehingga database yang akan dibuat perlu ditentukan durasi penyimpanannya yang paling ideal, hal ini bertujuan untuk menjaga tingkat akurasi data.

f. Melacak perubahan dari dimensi secara perlahan (Determine the need to track slowly changing dimensions)

Mengamati perubahan yang terjadi dari tabel dimensi dapat dilakukan dengan tiga tipe cara yaitu:

1. Tipe 1, dimana sebuah dimension attribute yang berubah dihapus dan digantikan dengan yang nilai baru (overwritten), misalnya ada perubahan nama produk tertentu, maka nama produk yang berubah langsung ditulis ulang.

2. Tipe 2, dimana sebuah dimension attribute yang berubah menyebabkan dimensions record baru tercipta. misalnya ada perubahan nama produk tertentu, maka nama produk yang berubah akan membentuk record baru tapi dengan kd_produk yang sama.
3. Tipe 3, dimana sebuah dimension attribute yang berubah menyebabkan sebuah attribute alternatif tercipta sehingga nilai attribute dengan nilai lama dan attribute dengan baru dapat diakses secara bersamaan di dalam dimension record data yang sama.

Dari ketiga tipe cara tersebut, perancangan dalam penelitian ini memilih perubahan pertama pada dimensi yang akan melakukan overwrite atas perubahan yang terjadi.

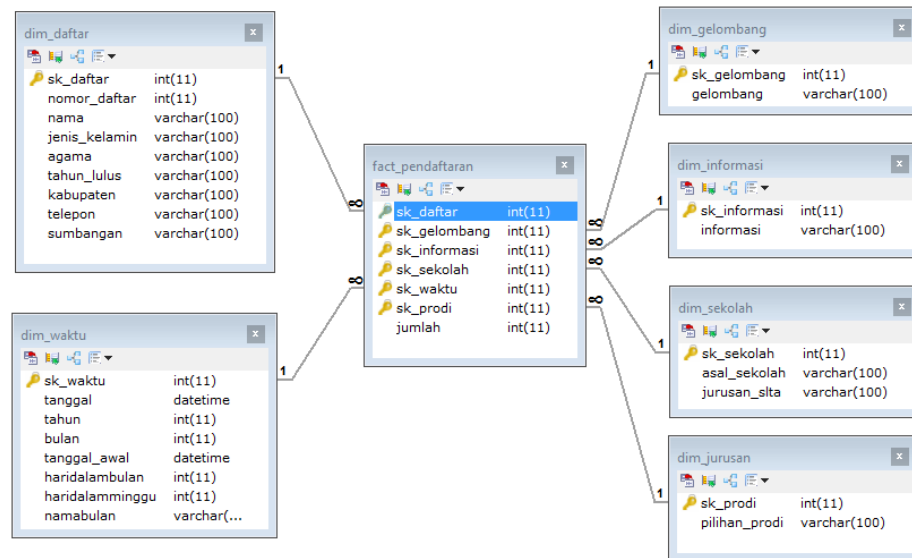
- g. Memilih prioritas query dan mode query (Decide the query priorities and the query modes)

Di bagian ini akan dijelaskan mengenai proses ETL. Sumber data yang digunakan dalam bentuk excel dimana beberapa record masih tidak standard dan memerlukan penyeragaman/cleaning data untuk kemudian dimasukkan ke dalam data warehouse.

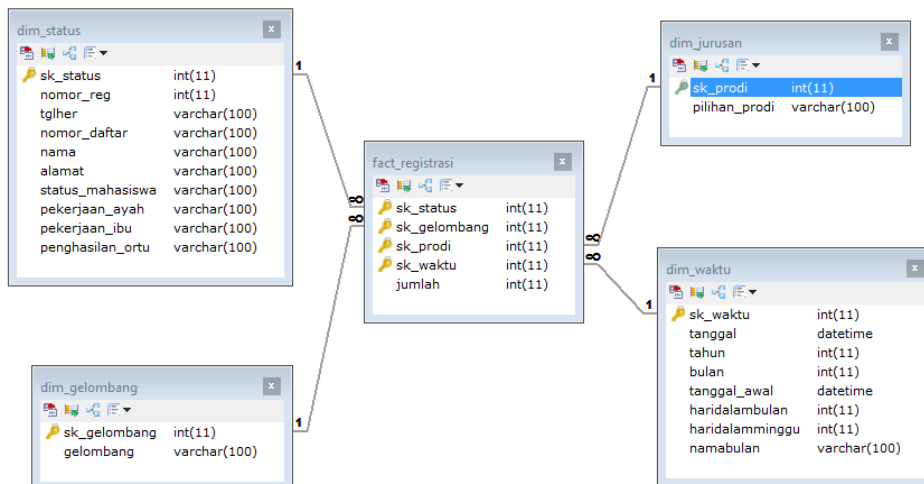
Dimensional Data Modelling

Perancangan dalam data warehouse penerimaan calon mahasiswa baru ini menggunakan Star Schema, bentuk ini dipilih karena tabel dimensinya tidak mengandung ringkasan atau tidak memiliki perbedaan tingkat ukuran sehingga tidak memerlukan tabel sub dimensi. Skema ini merupakan skema yang mudah dipahami daripada skema yang lain. Dengan menggunakan Star Schema, performa proses querynya menjadi lebih dan waktu pemrosesan menjadi lebih cepat, secara garis besar Star Schema terdiri dari satu tabel fakta dan beberapa tabel dimensi.

Berikut gambar star schema yang akan digunakan untuk implementasi OLAP data pendaftaran dan registrasi calon mahasiswa baru.



Gambar 5. Star Schema untuk data warehouse Pendaftaran



Gambar 6. Star Schema untuk data warehouse registrasi

Perancangan Database

Pada proses implementasi ETL aplikasi yang digunakan adalah Pentaho Data Intergration, Pada perancangan proses ETL ini langkah-langkah yang dilakukan adalah membuat beberapa koneksi database:

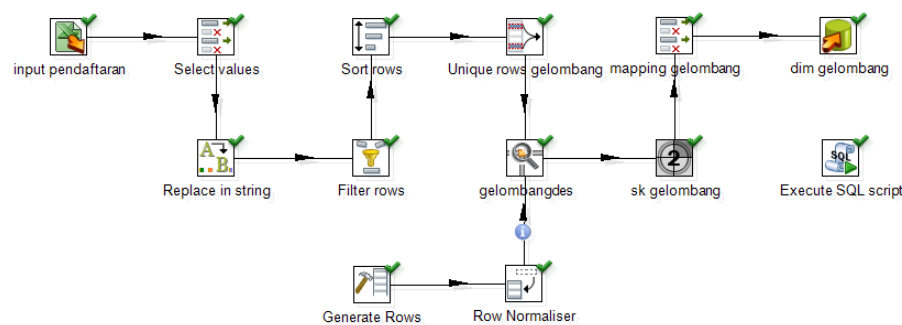
Implementasi ETL (*Extract, Transformation & Load*)

Pada tahap ini membahas tentang menggabungkan data dari sumber internal dengan proses *Extraction Transformation Load* (ETL), yaitu proses query dari sumber data, lalu dilakukan perubahan, pembersihan menjadi data yang konsisten,

data yang sudah di transformasi selanjutnya disimpan ke dalam tabel dimensi data warehouse. Tools yang digunakan untuk proses ETL ini adalah Pentaho Data Integration (PDI). Pada tahap ETL *design & development* ini terbagi atas dua tahap, yakni *dimension table staging* dan *fact table staging*.

Membentuk Dimensi Gelombang

Proses ETL dimulai dari pembacaan data sumber transaksional pendaftaran yang berasal dari Ms-excel. Langkah proses ETL dimensi Gelombang sebagai berikut:



Gambar 3.18. Proses ETL dim Gelombang

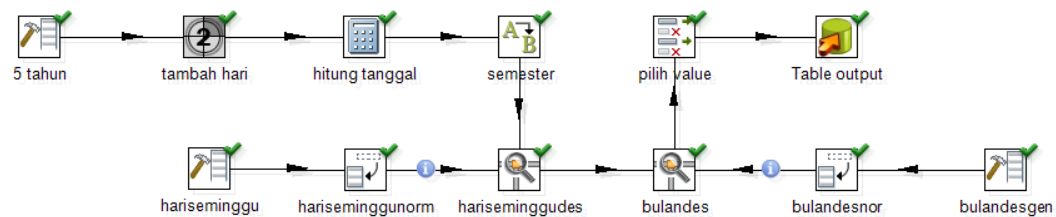
Gambar dibawah ini menunjukkan hasil eksekusi (*Execution Result*) dari dimensi gelombang. Total 4510 data berhasil dibaca dengan baik dalam waktu sekitar 6 (enam) detik.

Execution Results												
Execution History Logging Step Metrics Performance Graph												
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	sk gelombang	0	12	12	0	0	0	0	0	Finished	6.4s	2
2	mapping gelombang	0	12	12	0	0	0	0	0	Finished	6.4s	2
3	dim gelombang	0	12	12	0	12	0	0	0	Finished	6.5s	2
4	input pendaftaran	0	0	4510	4510	0	0	0	0	Finished	6.3s	713
5	Select values	0	4510	4510	0	0	0	0	0	Finished	6.3s	713
6	Filter rows	0	4510	4510	0	0	0	0	0	Finished	6.3s	712
7	Sort rows	0	4510	4510	0	0	0	0	0	Finished	6.4s	708
8	Unique rows gelombang	0	4510	12	0	0	0	0	0	Finished	6.4s	707
9	Replace in string	0	4510	4510	0	0	0	0	0	Finished	6.3s	712
10	Row Normaliser	0	1	12	0	0	0	0	0	Finished	1.0s	12
11	gelombangdes	0	24	12	0	0	0	0	0	Finished	6.4s	4
12	Generate Rows	0	0	1	0	0	0	0	0	Finished	0.6s	2
13	Execute SQL script	0	0	1	0	0	0	0	0	Finished	1.4s	1

Gambar 3.19. Proses Execution Results ke Tabel dim Gelombang

Membentuk Dimensi Waktu

Proses ETL dimulai dari pembacaan data sumber transaksional pendaftaran yang berasal dari Ms-excel. Langkah proses ETL dimensi Waktu sebagai berikut:

Gambar 3.20. Proses *ETL* dim Waktu

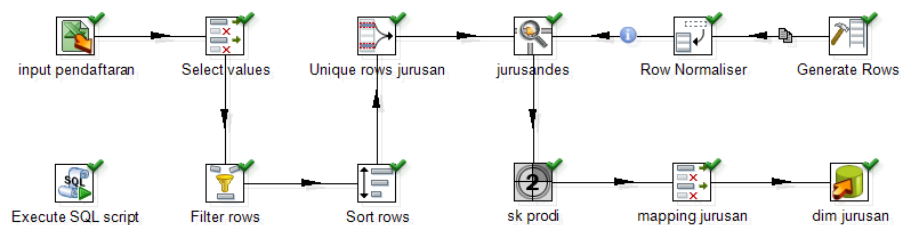
Gambar dibawah ini menunjukkan hasil eksekusi (*Execution Result*) dari dimensi waktu. Total 71826 data berhasil dibaca dengan baik dalam waktu sekitar 8 (delapan) detik.

Execution Results												
Execution History Logging Step Metrics Performance Graph												
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	tambah hari	0	1826	1826	0	0	0	0	0	Finished	0.1s	19,426
2	hitung tanggal	0	1826	1826	0	0	0	0	0	Finished	0.1s	18,444
3	5 tahun	0	0	1826	0	0	0	0	0	Finished	0.0s	107,412
4	semester	0	1826	1826	0	0	0	0	0	Finished	0.1s	17,558
5	hariseminggudes	0	71826	1826	0	0	0	0	0	Finished	0.4s	176,477
6	hariseminggu	0	0	10000	0	0	0	0	0	Finished	0.1s	112,360
7	hariseminggunorm	0	10000	70000	0	0	0	0	0	Finished	0.2s	339,806
8	bulandesnor	0	1	12	0	0	0	0	0	Finished	0.1s	132
9	bulandes	0	1838	1826	0	0	0	0	0	Finished	0.4s	4,483
10	bulandesgen	0	0	1	0	0	0	0	0	Finished	0.1s	11
11	pilih value	0	1826	1826	0	0	0	0	0	Finished	0.4s	4,411
12	Table output	0	1826	1826	0	1826	0	0	0	Finished	0.8s	2,314

Gambar 3.21. Proses *Execution Results* ke Tabel dim Waktu

Membentuk Dimensi Jurusan

Proses ETL dimulai dari pembacaan data sumber transaksional pendaftaran yang berasal dari Ms-excel. Langkah proses *ETL* dimensi Jurusan sebagai berikut:

Gambar 3.22. Proses *ETL* dim Jurusan

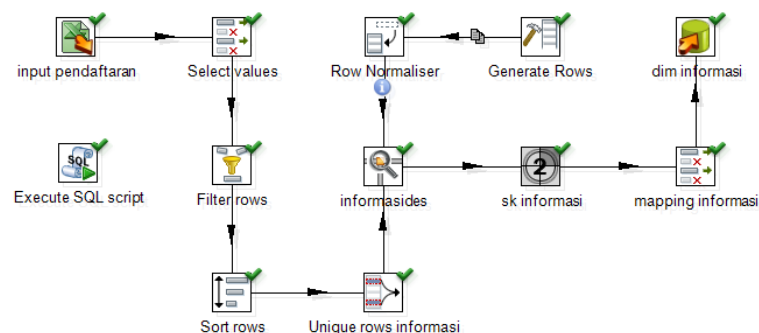
Gambar dibawah ini menunjukkan hasil eksekusi (*Execution Result*) dari dimensi Jurusan. Total 4510 data berhasil dibaca dengan baik dalam waktu sekitar 4 (empat) detik.

Execution History Logging Step Metrics Performance Graph												
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	Filter rows	0	4510	4505	0	0	0	0	0	Finished	3.8s	1,192
2	Sort rows	0	4505	4505	0	0	0	0	0	Finished	3.8s	1,180
3	Generate Rows	0	0	1	0	0	0	0	0	Finished	0.0s	143
4	Row Normaliser	0	1	3	0	0	0	0	0	Finished	0.5s	6
5	jurusandes	0	6	3	0	0	0	0	0	Finished	3.9s	2
6	Select values	0	4510	4510	0	0	0	0	0	Finished	3.8s	1,192
7	Unique rows jurusan	0	4505	3	0	0	0	0	0	Finished	3.8s	1,174
8	sk prodi	0	3	3	0	0	0	0	0	Finished	3.9s	1
9	mapping jurusan	0	3	3	0	0	0	0	0	Finished	3.9s	1
10	dim jurusan	0	3	3	0	3	0	0	0	Finished	4.0s	1
11	input pendaftaran	0	0	4510	4510	0	0	0	0	Finished	3.8s	1,194
12	Execute SQL script	0	0	1	0	0	0	0	0	Finished	0.0s	125

Gambar 3.23. Proses *Execution Results* ke Tabel dim Jurusan

Membentuk Dimensi Informasi

Proses ETL dimulai dari pembacaan data sumber transaksional pendaftaran yang berasal dari Ms-excel. Langkah proses *ETL* dimensi Informasi sebagai berikut:



Gambar 3.24. Proses *ETL* dim Informasi

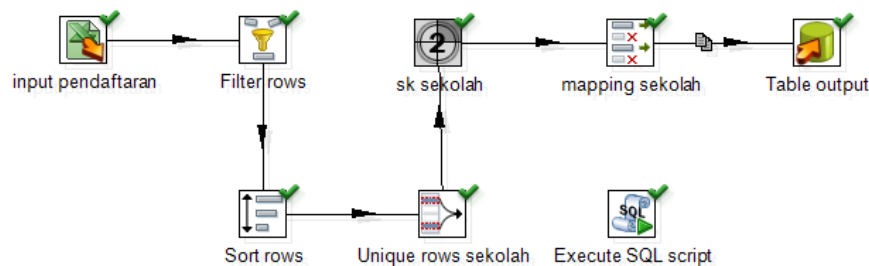
Gambar dibawah ini menunjukkan hasil eksekusi (*Execution Result*) dari dimensi Informasi. Total 4510 data berhasil dibaca dengan baik dalam waktu sekitar 5.5 (Lima koma lima) detik.

Execution Results												
Execution History Logging Step Metrics Performance Graph												
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	Filter rows	0	4510	4503	0	0	0	0	0	Finished	5.4s	840
2	Sort rows	0	4503	4503	0	0	0	0	0	Finished	5.4s	834
3	Generate Rows	0	0	1	0	0	0	0	0	Finished	0.0s	100
4	Row Normaliser	0	1	6	0	0	0	0	0	Finished	0.4s	14
5	informasides	0	12	6	0	0	0	0	0	Finished	5.4s	2
6	Unique rows informasi	0	4503	6	0	0	0	0	0	Finished	5.4s	832
7	sk informasi	0	6	6	0	0	0	0	0	Finished	5.4s	1
8	input pendaftaran	0	0	4510	4510	0	0	0	0	Finished	5.4s	842
9	Select values	0	4510	4510	0	0	0	0	0	Finished	5.4s	841
10	mapping informasi	0	6	6	0	0	0	0	0	Finished	5.5s	1
11	dim informasi	0	6	6	0	6	0	0	0	Finished	5.5s	1
12	Execute SQL script	0	0	1	0	0	0	0	0	Finished	0.0s	71

Gambar 3.25. Proses *Execution Results* ke Tabel dim Informasi

Membentuk Dimensi Sekolah

Proses ETL dimulai dari pembacaan data sumber transaksional pendaftaran yang berasal dari Ms-excel. Langkah proses *ETL* dimensi Sekolah sebagai berikut:

Gambar 3.26. Proses *ETL* dim Sekolah

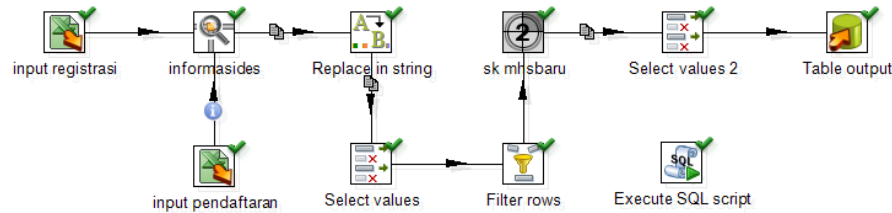
Hasil eksekusi (*Execution Result*) dari dimensi Sekolah. Total 4510 data berhasil dibaca dengan baik dalam waktu sekitar 11 (sebelas) detik.

Execution Results												
Execution History Logging Step Metrics Performance Graph												
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	sk sekolah	0	914	914	0	0	0	0	0	Finished	11.1s	82
2	mapping sekolah	0	914	914	0	0	0	0	0	Finished	11.1s	82
3	Table output	0	914	914	0	914	0	0	0	Finished	11.4s	80
4	Unique rows sekolah	0	4490	914	0	0	0	0	0	Finished	11.1s	405
5	Filter rows	0	4510	4490	0	0	0	0	0	Finished	11.1s	408
6	Sort rows	0	4490	4490	0	0	0	0	0	Finished	11.1s	405
7	input pendaftaran	0	0	4510	4510	0	0	0	0	Finished	11.1s	408
8	Execute SQL script	0	0	1	0	0	0	0	0	Finished	0.1s	11

Gambar 3.27. Proses *Execution Results* ke Tabel dim Sekolah

Membentuk Dimensi Mhsbaru

Proses ETL dimulai dari pembacaan data sumber transaksional pendaftaran yang berasal dari Ms-excel. Langkah proses *ETL* dimensi Mahasiswa Baru sebagai berikut:



Gambar 3.28. Proses *ETL* dim Mhsbaru

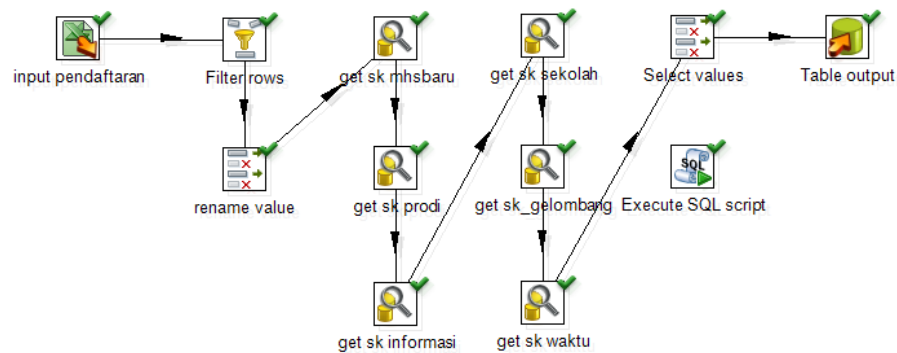
Gambar dibawah ini menunjukkan hasil eksekusi (*Execution Result*) dari dimensi Mhsbaru. Total 4510 data berhasil dibaca dengan baik dalam waktu sekitar 17 (tujuh belas) detik.

Execution Results												
Execution History Logging Step Metrics Performance Graph												
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	Select values 2	0	3450	3450	0	0	0	0	0	Finished	16.6s	208
2	Table output	0	3450	3450	0	3450	0	0	0	Finished	17.2s	200
3	input pendaftaran	0	0	4510	4510	0	0	0	0	Finished	16.3s	276
4	informasides	0	7960	3450	0	0	0	0	0	Finished	16.5s	481
5	input registrasi	0	0	3450	3450	0	0	0	0	Finished	12.0s	288
6	Replace in string	0	3450	3450	0	0	0	0	0	Finished	16.5s	208
7	Select values	0	3450	3450	0	0	0	0	0	Finished	16.5s	208
8	Filter rows	0	3450	3450	0	0	0	0	0	Finished	16.6s	208
9	sk mhsbaru	0	3450	3450	0	0	0	0	0	Finished	16.6s	208
10	Execute SQL script	0	0	1	0	0	0	0	0	Finished	0.5s	2

Gambar 3.29. Proses *Execution Results* ke Tabel dim Mhsbaru

Fact Table Staging Pendaftaran

Tahap selanjutnya adalah table fakta, yaitu tabel fakta Pendaftaran, yang sesuai dimensional modeling dibentuk dari tabel-tabel dimensi dengan skema bintang (*star schema*). Pada gambar 3.30. menjelaskan proses ETL untuk membentuk tabel fakta. Menjelaskan proses yang dimulai dari Extract, Transform sampai Loading terhadap tabel data transaksional dan tabel dimensi sehingga membentuk tabel fakta pendaftaran ke *data warehouse*.



Gambar 3.30. Proses ETL fakta pendaftaran

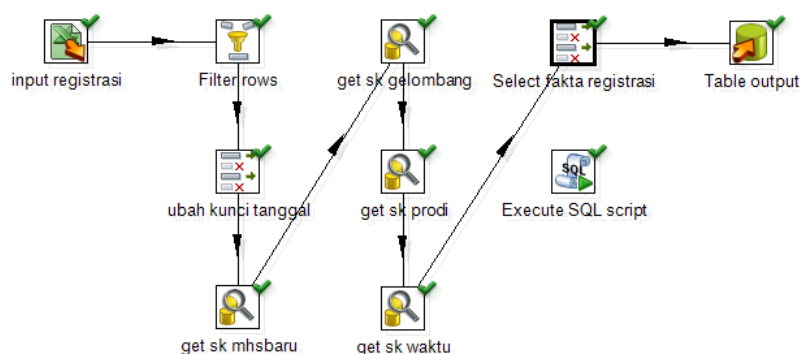
Gambar dibawah ini menunjukkan hasil eksekusi (*Execution Result*) dari fakta pendaftaran. Total 4510 data berhasil dibaca dengan baik dalam waktu sekitar 32 (tiga puluh dua) detik.

Execution Results												
Execution History Logging Step Metrics Performance Graph												
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	get sk informasi	0	4510	4510	4503	0	0	0	0	Finished	30.5s	148
2	get sk sekolah	0	4510	4510	4507	0	0	0	0	Finished	30.5s	148
3	Select values	0	4510	4510	0	0	0	0	0	Finished	30.9s	146
4	Table output	0	4510	4510	0	4510	0	0	0	Finished	31.6s	142
5	get sk prodi	0	4510	4510	4505	0	0	0	0	Finished	30.5s	148
6	get sk gelombang	0	4510	4510	1398	0	0	0	0	Finished	30.5s	148
7	get sk waktu	0	4510	4510	4510	0	0	0	0	Finished	30.8s	146
8	get sk mhsbaru	0	4510	4510	3533	0	0	0	0	Finished	30.5s	148
9	Filter rows	0	4510	4510	0	0	0	0	0	Finished	8.0s	566
10	rename value	0	4510	4510	0	0	0	0	0	Finished	8.0s	566
11	input pendaftaran	0	0	4510	4510	0	0	0	0	Finished	8.0s	567
12	Execute SQL script	0	0	1	0	0	0	0	0	Finished	0.5s	2

Gambar 3.31. Proses *Execution Results* ke Tabel fakta pendaftaran

Fact Table Staging Registrasi

Tahap selanjutnya adalah table fakta, yaitu tabel fakta Registrasi, yang sesuai dimensional modeling dibentuk dari tabel-tabel dimensi dengan skema bintang (*star schema*). Pada gambar 3.32. menjelaskan proses ETL untuk membentuk tabel fakta. Menjelaskan proses yang dimulai dari Extract, Transform sampai Loading terhadap tabel data transaksional dan tabel dimensi sehingga membentuk tabel fakta registrasi ke *data warehouse*.



Gambar 3.32. Proses ETL fakta registrasi

Gambar dibawah ini menunjukkan hasil eksekusi (*Execution Result*) dari fakta registrasi. Total 3450 data berhasil dibaca dengan baik dalam waktu sekitar 20 (dua puluh) detik.

Execution Results												
Execution History Logging Step Metrics Performance Graph												
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	Select fakta registrasi	0	3450	3450	0	0	0	0	0	Finished	19.2s	180
2	Table output	0	3450	3450	0	3450	0	0	0	Finished	19.8s	174
3	get sk gelombang	0	3450	3450	455	0	0	0	0	Finished	18.9s	183
4	get sk prodi	0	3450	3450	3441	0	0	0	0	Finished	18.9s	183
5	get sk waktu	0	3450	3450	3450	0	0	0	0	Finished	19.2s	180
6	Filter rows	0	3450	3450	0	0	0	0	0	Finished	6.2s	560
7	ubah kunci tanggal	0	3450	3450	0	0	0	0	0	Finished	6.2s	559
8	get sk mhsbaru	0	3450	3450	3184	0	0	0	0	Finished	18.9s	183
9	input registrasi	0	0	3450	3450	0	0	0	0	Finished	6.2s	560
10	Execute SQL script	0	0	1	0	0	0	0	0	Finished	0.9s	1

Gambar 3.33. Proses *Execution Results* ke Tabel fakta registrasi

Penjadwalan Proses Otomatisasi ETL dan Job

Setelah proses transformasi data menjadi bentuk tabel dimensi atau bentuk tabel fakta. Semua transformasi tersebut masih dicontohkan untuk dijalankan secara manual. Untuk membuat transformasi menjadi otomatis Pentaho menamakan otomatisasi ETL dengan job. Fungsinya antara lain menentukan Desain job dan melakukan *Setting job scheduling*.

Analisis Hasil Data Warehouse

Analisis pada data warehouse pada penelitian ini dilakukan dengan Pentaho Report Designer adalah tools untuk merancang laporan sesuai dengan business requirement. Selain relative mudah penggunaannya, Pentaho Report Designer juga dapat menghasilkan outputan ke banyak format (.pdf, .xls, .html, dll). Adapun contoh report yang telah dibuat adalah sebagai berikut:

- a. Bagaimana laporan calon mahasiswa baru per-jurusan dan per-informasi

STMIK Amikom Purwokerto

Rekap Laporan Penerimaan Mahasiswa Baru
Periode 2011 - 2015

Per Jurusan Per Informasi

PILIHAN PRODITI (SORE)				
NOMOR DAFTAR	TGLDAFTAR	NAMA	JURUSAN SLTA	INFORMASI
110001	2011-01-13	ANUGRAH ANANG PRASETYO	AKUNTANSI	BROSUR
PILIHAN PRODISI				
NOMOR DAFTAR	TGLDAFTAR	NAMA	JURUSAN SLTA	INFORMASI
110002	2011-01-13	WISNU BROTO B.U	IPS	LAINNYA
110003	2011-01-17	FITRIANINGSIH	AKOMODASI PERHOTELAN	BROSUR
PILIHAN PRODITI				
NOMOR DAFTAR	TGLDAFTAR	NAMA	JURUSAN SLTA	INFORMASI
110004	2011-01-17	ARIE SETIADI	IPS	BROSUR
110005	2011-01-18	RIZKY ASSYARUFY	IPS	BROSUR

Gambar 3.40. laporan per-jurusan berdasarkan per-informasi

Pengujian Sistem

Hasil uji sistem *data warehouse* penerimaan mahasiswa baru berupa kuisioner data penilaian fungsi yang digunakan dalam *reporting data warehouse*, didapatkan data sebagai berikut:

- a. Pengujian Karakteristik *Functionality*

$$\sum X_1 = 101,2 \text{ (Jumlah rata-rata skor pertanyaan)}$$

$$n = 11 \text{ (Jumlah Responden)}$$

$$\text{Skor} = 99,6 / 11$$

$$= 9,05$$

- b. Pengujian *Reability*

$$\sum X_1 = 96,17 \text{ (Jumlah rata-rata skor pertanyaan)}$$

$$n = 21 \text{ (Jumlah Responden)}$$

$$\text{Skor} = 96,17 / 21$$

$$= 4,58$$

c. Pengujian *Usability*

$\sum X_1 = 103,4$ (Jumlah rata-rata skor pertanyaan)

$n = 23$ (Jumlah Responden)

Skor = $103,4 / 11$

= 4,50

d. Pengujian *Efficiency*

$\sum X_1 = 95,8$ (Jumlah rata-rata skor pertanyaan)

$n = 23$ (Jumlah Responden)

Skor = $95,8 / 11$

= 4,17

e. Pengujian *Maintainability*

$\sum X_1 = 92$ (Jumlah rata-rata skor pertanyaan lampiran 7)

$n = 23$ (Jumlah Responden)

Skor = $92 / 11$

= 4

f. Pengujian *Portability*

$\sum X_1 = 87,2$ (Jumlah rata-rata skor pertanyaan lampiran 8)

$n = 23$ (Jumlah Responden)

Skor = $96,17 / 11$

= 3,79

KESIMPULAN DAN SARAN

A. Kesimpulan

Kesimpulan yang dapat dikemukakan antara lain sebagai berikut.

1. Dengan dibangunnya data warehouse Pemasaran di STMIK Amikom Purwokerto, maka penelusuran informasi yang berkaitan dengan penerimaan mahasiswa baru dapat dilakukan dengan mudah dan lebih fleksibel.
2. Informasi yang disajikan dari data warehouse Pemasaran dapat dipergunakan untuk hal-hal sebagai berikut ini.

3. Membantu Pimpinan STMIK Amikom Purwokerto dalam proses evaluasi dan perencanaan promosi ke daerah-daerah dan sekolah-sekolah asal calon mahasiswa yang potensial bagi STMIK Amikom Purwokerto.
4. Membantu Pimpinan STMIK Amikom Purwokerto dalam proses evaluasi dan perencanaan penggunaan media promosi yang digunakan dimasing-masing daerah.
5. Membantu Pimpinan STMIK Amikom Purwokerto saat mencermati trend pilihan calon mahasiswa baru pada tiap-tiap program studi di setiap periode sehingga Pimpinan STMIK Amikom Purwokerto dapat dengan segera melakukan langkah-langkah preventive terhadap trend yang sedang berjalan.
6. Membantu Pimpinan STMIK Amikom Purwokerto dalam proses evaluasi kinerja bagian pemasaran.

B. Saran

Berikut ini adalah saran yang ditujukan untuk perbaikan dimasa yang akan datang.

1. Struktur basis data akademik pada STMIK Amikom Purwokerto sebaiknya dirancang menggunakan kaidah normalisasi sehingga lebih efisien dan dapat dirancang menjadi lebih lengkap sesuai dengan kebutuhan.
2. Hendaknya dilakukan validasi dan standarisasi terhadap masukan data yang di-entry pada aplikasi-aplikasi yang dibangun di STMIK Amikom Purwokerto sehingga penelusuran informasi menjadi lebih mudah dan terhindar dari masalah inkonsistensi data.
3. STMIK Amikom Purwokerto hendaknya dapat melengkapi kekurang-kekurangan pada data akademiknya sehingga teknik data mining yang dibangun dapat menghasilkan rule- rule yang ideal dan dapat diterapkan pada seluruh program studi yang ada di STMIK Amikom Purwokerto.

DAFTAR PUSTAKA

- Goldstein, P.J., 2005, *Academic Analytics: The Uses of Management Information and Technology in Higher Education*, ECAR Key Findings, EDUCAUSE
- Han, Jiawei, Micheline Kamber, 2006, *Data Mining: Concepts and Techniques*. 2nd edition, San Fransisco: Morgan Kaufmann Publishers.
- Inmon, W. H., 2002, *Building the Data warehouse*, 4rd edition, John Wiley & Sons, New York, NY: United States Of America
- JRP, Mulyana, 2014, *Pentaho: Solusi Open Source untuk Membangun Data Warehouse*, Penerbit Andi, Yogyakarta.
- Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., dan Becker, B., 2002, *The Data warehouse Lifecycle Toolkit*, second edition, Wiley Publishing, Indianapolis, IN: United States Of America
- Karyono, Giat, 2010, *Pengembangan data warehouse dan on-line analytical processing (OLAP) untuk penemuan informasi dan analisis data*, Jurnal Telematika, Vol. 5, no. 2, Indonesia
- Kimball, R., dan Ross, M., 16 Maret 2015, *The Kimball Bus Architecture and the Corporate Information Factory: What are the fundamental differences?*, <http://www.informationweek.com/news/showArticle.jhtml?articleID=17800088>
- Nabuzin, 2012, *Data warehouse from Architecture to Implementation*, Reading, MA: Addison-Wesley, Boston, United States Of America
- Nazir, M., 2005, *Metode Penelitian*, Ciawi, Penerbit Ghalia Indonesia
- P., Hendro Poerbo, 04 Maret 2015, *Purwarupa Data warehouse Pada Sistem Informasi Manajemen Perguruan Tinggi Studi Kasus : Stikom Surabaya*, <http://digilib.its.ac.id/public/ITS-Master-14827-paperpdf.pdf>
- Pressman, R. S., 2012, *Rekayasa Perangkat Lunak*. Jakarta, Penerbit Andi. Yogyakarta.
- Suryanto, Wahyu Dwi, 03 Maret 2015, *Pengembangan Data warehouse Dan Aplikasi Olap Data Tracer Study Alumni IPB Berbasis Web Menggunakan Microsoft Business Intelligence*, <http://wdwisuryanto.staff.ipb.ac.id/files/2012/01/MakalahPraseminar1.pdf>
- S, Restia Rezalini P., Anggraeni, Wiwik., Wibowo, Radityo Prasetyanto, 04 Maret 2015, *Perancangan Dan Pembuatan Data warehouse Untuk Kebutuhan Sistem Pendukung Keputusan di Bidang Akademik Pada Jurusan Sistem Informasi, ITS, Surabaya*, <http://digilib.its.ac.id/public/ITS-Undergraduate-14414-paperpdf.pdf>

- Sumpala, Andi Tenri, 25 Maret 2015, Perhitungan indikator kinerja rumah sakit menggunakan data warehouse di rumah sakit, <http://www.sciencedirect.com/science/article/pii/S1067502703000884>
- Wikramanayake, G.N., Goonetillake J.S., 2006, Managing Very Large Databases and Data Warehousing, Sri Lankan Journal of Librarianship and Information Management, Vol. 2, No.1, pp. 22-29, Pakistan
- Yudi Wibisono, 2012, Modul Pengantar Pentaho Kettle, Ilmu Komputer UPI
- Yohanlis, Stephanie, 22 Maret 2015, Perancangan Data warehouse pada Profil Mahasiswa di Universitas, <http://ejournal.ukrida.ac.id/ojs/index.php/TK/article/download/846/1113>