Rancangan dan Pembuatan Data Warehouse untuk Kebutuhan Sistem Pendukung Keputusan

(Studi Kasus: Departemen Marketing dan Komunikasi Universitas Dr Soetomo Surabaya)

Slamet Kacung[1], Lambang Probo S[2]

Abstract

Information about the distribution of data and supporting students is needed by marketing and communications department (markom) Dr. Soetomo University (Unitomo) of Surabaya. On the other hand, the data distribution information markom students rely on the IT Department. To overcome these problems need to developed data warehouse applications in order to support environmentally markkom routine activities and faculty in promoting Unitomo to the wider community. The data warehouse is a data set that is subject-oriented, integrated, time variant, and nonvolatile which helps management in decision making process.

The method of data warehouse started from collecting the data needed to build a data warehouse, the next process is the star schema created, extraction and data transformation (ETL). Data extraction is the process of selecting the data to be inserted into the data warehouse. Once the transformation is done, the data is inserted into the data warehouse. Existing data in the data warehouse with OLAP processed (Online Analytical Processing) to produce information.

The purpose of a data warehouse is to provide early information to markom in conducting an analysis of the distribution of the data's student performance is measured and in accordance with the scheme or plan that you want to work in the promotion for the next year.

***Keyword* :** Data Warehouse, OLAP, ETL, Star Schema.

Pendahuluan

Kinerja markom merupakan indikator dari dari jumlah mahasiswa yang didapat dari setiap perguruan tinggi swasta (PTS), PTS dapat dikatakan dalam kondisi buruk apabila mengalami penurunan jumlah mahasiswa. Salah satu yang bias dilakukan oleh markom untuk menjaga agar jumlah student body tetap stabil adalah dengan menganalisis dan mengetahui faktor-faktor yang dapat menjaga kinerja promosi itu sendiri berada dalam posisi baik. Analisis dapat dilakukan dengan mengumpulkan data sebaran mahasiswa yang bersifat historis atau lampau. Semakin banyak data yang dimiliki, maka semakin banyak pula hasil analisis yang dihasilkan. Namun, untuk memenuhi hasil analisis yang baik dibutuhkan suatu teknologi yang dapat membantu pihak markom dalam memahami analisis yang akan dilakukan.

Data warehouse adalah sebuah database yang secara khusus didesain dengan struktur untuk melakukan query dan analisis (Nolan & Huguelet, 2000). Data

[1] Fakultas Teknik, Program Studi Teknik Informatika Universitas Dr. Soetomo, Telp : (031) 5944744, Fax:031-5938935, E-mail : slamet@unitomo.ac.id

[2] Fakultas Teknik, Program Studi Teknik Informatika Universitas Dr. Soetomo, Telp : (031) 5944744, Fax:031-5938935, E-mail : lapros@unitomo.ac.id

warehouse merupakan database komprehensif yang mendukung semua analisis keputusan yang diperlukan oleh suatu organisasi dengan menyediakan ringkasan dan rincian informasi (Turban, dkk, 2005). Data warehouse menyediakan suatu wadah untuk menampung data-data yang diperlukan untuk menganalisis suatu kondisi dalam organisasi dengan hanya mengambil data sesuai kebutuhan.

Mempunyai basis data relational memang tidak berarti mendukung pembuatan keputusan. Waktu respon yang lambat, sistem yang kurang fleksibel dan penggunaan sumber daya jaringan yang berlebihan adalah cirri aplikasi analitis yang dibangun diatas dasar teknologi basis data relational seperti sekarang yang dimiliki Universitas Dr. Soetomo (Unitomo) Surabaya. Basis data relational tidak ditujukan untuk menyediakan fungsi yang mampu melakukan sintesa data, analisis dan konsolidasi yang kemudian dikenal sebagai analisis data multidimensi (OLAP).

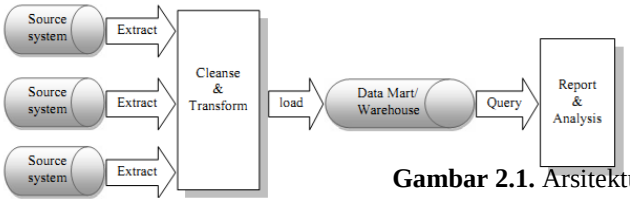
Mengatasi hal tersebut, markom unitomo membutuhkan sebuah aplikasi OLAP untuk analisis data mahasiswa yang fleksibel, cepat merespon permintaan informasi dari manajemen yang dapat di kontrol secara terpusat sehingga kebutuhan pemeliharaan dapat ditekan seminimal mungkin.

Tinjauan Pustaka

Data Warehouse

Menurut Singh (1999), data warehouse bukanlah sebuah produk tunggal tetapi sebuah lingkungan yang terdiri dari banyak produk dari banyak vendor, seperti pada Gambar 2.1, yang meliputi :

1. Sistem sumber, tempat data diambil
2. Tools untuk mengekstrak data ke data warehouse (transmission tool)
3. Basis data data warehouse
4. Desktop query dan reporting tools untuk mendukung keputusan



Gambar 2.1. Arsitektur Data Warehouse

Karakteristik Data Warehouse

Terdapat empat karakteristik data warehouse yang diperlukan (Inmon,2002) :

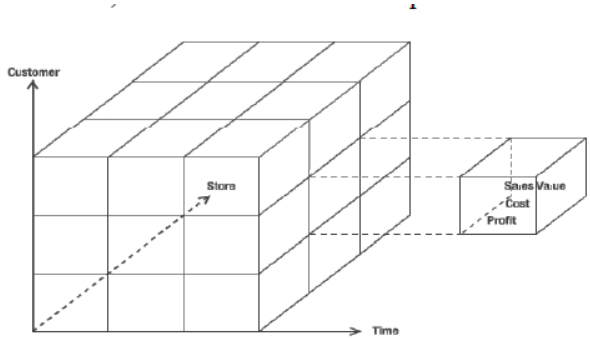
1. Subyek Oriented
   Data yang dimasukkan dalam data warehouse adalah data yang benar-benar diperlukan dalam proses pengambilan dan menganalisis data tertentu.
2. integrated
   Data yang berasal dari berbagai sumber sudah disusun dengan konversi bersama sehingga satu nama data warehouse memiliki arti dan format yang sama pada semua database sumber.
3. Time Variant
   Data warehouse (historical data). Melakukan perbandingan dengan kebutuhan sistem operasional yang hampir semuannya data mutakhir.
4. Non-Volatile

   Sekali masuk kedalam data warehouse data tersebut tidak akan pernah di update atau dihapus

OLAP

Menurut Frosman (1997), definisi OLAP :

"OLAP is a process completed by people using tools facilitate the analysis of dimensionally structured business information. OLAP allows user to navigate data by drilling down to lower-level and pivoting. OLAP tools support some level of calculation against the data". Jadi dapat dikenal bahwa OLAP adalah sebuah proses pelengkap yang dilakukan dengan menggunakan tool tertentu untuk menganalisis informasi bisnis yang strukturnya dimensional, sehingga pengguna dapat menavigasi data dengan cara drill-down ke data yang lebih detail maupun drill-up ke data yang lebih ringkas, melakukan pivoting perhitungan tertentu terhadap data.



Gambar 2.2. Data Cube Multidimensi

Manfaat OLAP

Aplikasi OLAP yang sukses dapat meningkatkan produktifitas dari organisasi secara keseluruhan. Fleksibelitas OLAP memungkinkan pengguna OLAP menjadi mandiri, dimana para manajer tak lagi bergantung pada departemen teknologi informasi, dan dapat memodelkan masalah yang sebelumnya tak mungkin dapat dilakukan dengan sistem yang kurang fleksibel dan selalu menunggu support dari departement.

Departemen MIS juga dapat manfaat penggunaan software OLAP yang tepat, MIS dapat mengantarkan informasi ke pengguna secara lebih tepat dan cepat, dan karena sumber data OLAP adalah data warehouse maka ada kontrol terpusat atas data. Operasional jaringan juga menjadi lebih efisien, karena penggunaan OLAP mengurangi query dan drag dan lalu lintas jaringan atas OLTP maupun akses data warehouse secara langsung.

Sehingga pada akhirnya dengan menyediakan kemampuan untuk memodelkan masalah yang sebenarnya dan pengguna sumber daya manusia yang lebih efisien, maka OLAP memungkinkan organisasi secara keseluruhan untuk merespon secara cepat.

Tabel 2.1
Keuntungan menggunakan OLAP (Hyperion, 1998)

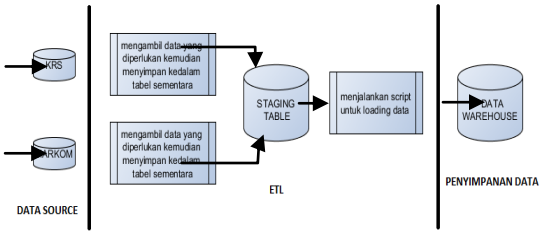
Keuntungan bagi pengguna	Keuntungan bagi Departemen MIS
✓ Pengguna tidak perlu mempelajari bahasa SQL ✓ Pengguna tidak perlu mengerti model basis data relasional ✓ Memperbaiki kinerja query dan skalabilitas sistem ✓ Meningkatkan kemampuan sistem dalam melakukan perhitungan	✓ Manajemen sistem yang lebih mudah ✓ Mengotomasikan pemeliharaan data ✓ Mengurangi beban data warehouse RDBMS ✓ Membebaskan beban data warehouse RDBMS ✓ Memungkinkan control terpusat atas data analitis

Rancangan Data Warehouse

Desain data warehouse diawali dengan penentuan level-of-granulity dari data, dilanjutkan dengan desain star-schema diagram dan low-level schema dari data, dan desain proses ECTL. Penentuan level-of-granulity dari data perolehan mahasiswa ditentukan sampai dengan detail asal sekolah, asal kota, per-fakultas, per-jurusan, per-tahun-akademik.

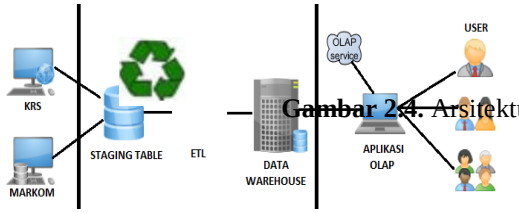
Perancangan Arsitektur

Tahap perancangan adalah tahapan kedua yang harus dilakukan untuk pembentukan data warehouse. Di dalamnya meliputi perancangan arsitektur logical maupun fisik dari data warehouse.



Gambar 2.3. Arsitektur Logical Data Warehouse

Perancangan arsitektur fisik pada data warehouse dapat dilihat pada gambar 2.4. berikut :



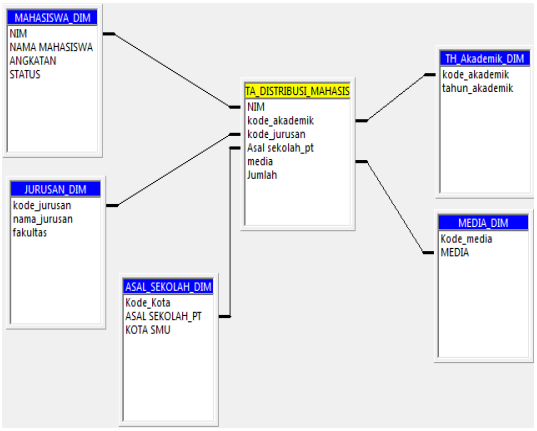
Gambar 2.4. Arsitektur Fisik Data Warehouse

Perancangan Star Schema

Pendekatan desain skema tabel yang dipergunakan untuk OLAP untuk membuat struktur informasi multidimensi yang cocok dengan kebutuhan adalah dengan pendekatan Star Schema. Karakteristik dari star schema adalah :

1. Pusat dari bintag adalah tabel fakta.
   - Tabel fakta berisi KPI yang relevan dari objek informasi dan waktu
   - KPI tersebut adalah attribute dari tabel fakta
   - Objek informasi dan waktu adalah elemen dari key
2. Titik-titik dari bintang adalah tabel dimensi
   - Tabel dimensi berisi data mengenai objek informasi atau waktu
   - Tabel fakta dan tabel dimensi dihubungkan melalui multipart primary pada tabel fakta
   - Setiap tabel dimensi dihubungkan secara langsung ke tabel fakta melalui sebuah kolom key.

Sesuai hasil analis kebutuhan sistem, dibutuhkan satu star schema yaitu star schema distribusi mahasiswa seperti yang terlihat pada gambar 2.5 seperti berikut :



Gambar 2.5 Star Schema Analisa Data Mahasiswa

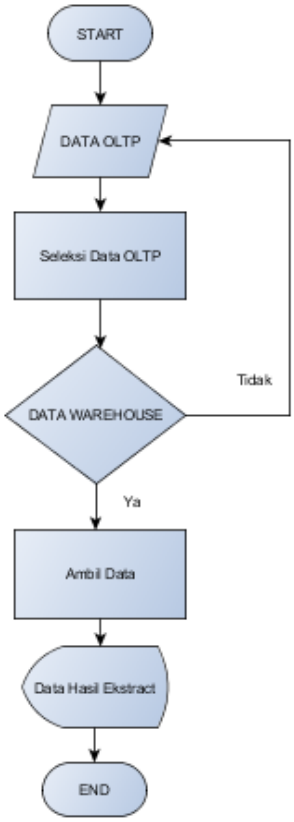
Proses ECTL

a. Proses Transformasi Data Extract

ETL adalah sebuah proses pengambilan data dari OLTP (sumber data), dengan ETL data operational dapat disimpan ke dalam sementara sebelum *(staging table)* data di *transform* ke data warehouse. *Extract*, *transform*, dan *load* (ETL) merupakan sebuah sistem yang dapat membaca data dari suatu data store, mengumpulkan, menyaring, mengolah dan menggabungkan data yang relevan dari berbagai sumber dan menyimpan ke data warehouse. Data store yang dibaca ETL

disebut *data source*, sedangkan *data store* yang disimpan ETL disebut target. Proses pengubahan data digunakan agar data sesuai dengan format dan kriteria, atau sebagai validasi data dari *source system*.

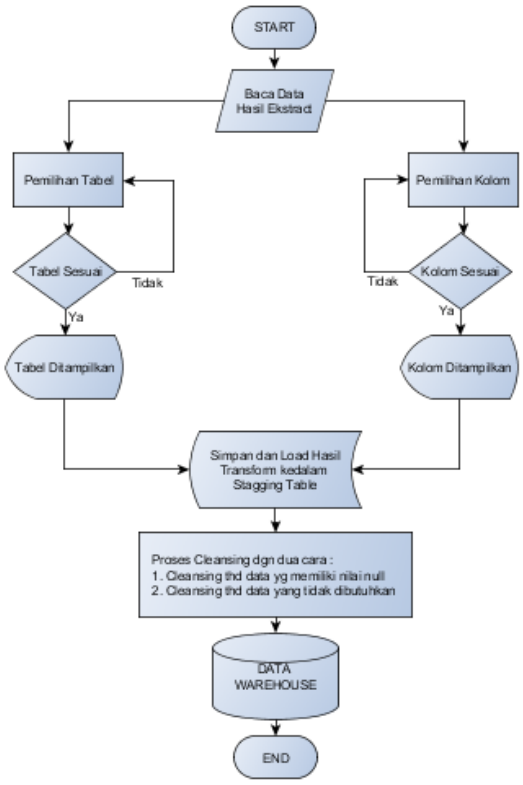
Proses pertama dari ETL adalah ekstraksi data, yaitu mengambil data dari sumber-sumber data yang berbeda dengan sistem yang terpisah yang menggunakan format data yang berbeda. Ekstraksi adalah mengubah data kedalam format yang berguna untuk proses transformasi. Berikut adalah proses ekstract pada data warehouse.



Gambar 2.6. *Flowchart* Proses *Extract*

b. Proses Cleansing

Proses pembersihan data *(cleansing)* harus dilakukan untuk menghindari ketidakkonsistenan terhadap data sebelum data di transformasi kedalam penyimpanan data pada data warehouse, karena data didalam data warehouse tidak dalam bentuk normalisasi. Sebelum proses pembersihan dilakukan terlebih dahulu membuat tabel sementara *(staging table)* untuk menampung data hasil dari proses *cleansing*,untuk lebih jelas mengenai proses pembersihan data akan dijelaskan dalam *flowchart* seperti pada gambar 2.7. Setelah proses *cleansing*, maka proses yang terakhir adalah proses *load*. Fase *load* merupakan tahapan yang berfungsi untuk memasukkan data kedalam target akhir. Pada proses ini data yang sudah dibaca, dibersihkan dan diubah formatnya akan disimpan pada data warehouse.



Gambar 2.7. *Flowchart* Proses *Cleansing*

c. Data Mining

Definisi data mining menurut Microsoft (2001) adalah "*Data mining is the process of discovering meaningful patterns and relationship that lie hidden within very large database*".

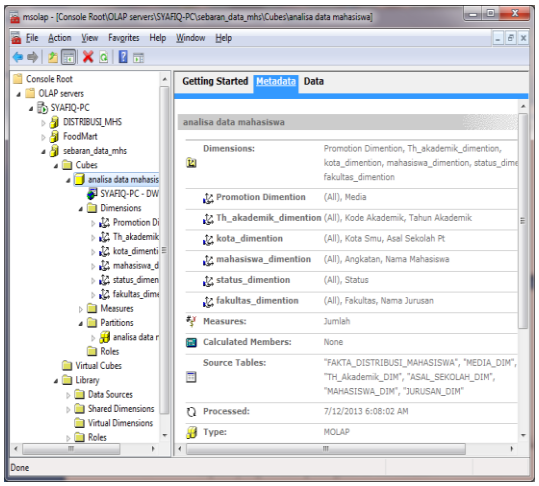
Menurut definisi tersebut dapat dikatakan bahwa data mining adalah proses untuk menemukan pola dan hubungan yang berarti yang tersembunyi dalam basis data yang amat besar. Jadi meskipun OLAP dan data mining keduanya adalah tool pendukung keputusan, tetapi masing-masing didesain untuk penggunaan yang berbeda.

Data mining sendiri bukanlah sebuah teknologi tunggal, melainkan adalah kumpulan dari tool yang digunakan untuk mengekstrak informasi dari data. Setiap tool punya kegunaan yang berbeda untuk masalah yang berbeda. Data mining tool dapat menjalankan query yang sangat sederhana, metode statitstik tradisional, sampai penggunaan seperangkat tool canggih yang dapat secara otomatis mencari hubungan/pola tersembunyi pada data warehouse

Hasil dan Pembahasan

a. Hasil Meta Data OLAP

Hasil dari file OLAP adalah berawal dari pembuatan kubus dengan diawali pembuatan beberapa dimensi regular sesuai dengan desain aweal, di lanjutkan dengan pembuatan kubus standart yaitu analisa data mahasiswa. pembuatan kubus pada SQL Server dapat dilakukan melalui OLAP Manager seperti yang terlihat pada gambar 2.8 berikut :



Gambar 2.8 Metadata Hasil OLAP

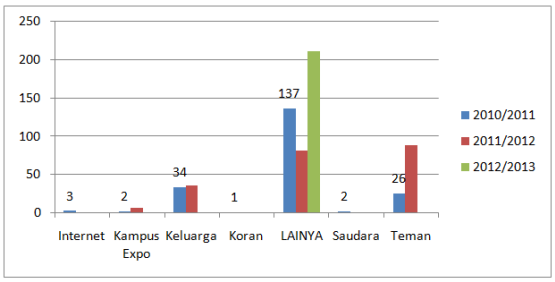
b. Hasil OLAP (Data Multideimensi)

analisa data mahasiswa

kota_dimention: All kota_dimention | mahasiswa_dimention: All mahasiswa_dimer
status_dimention: All status_dimention | fakultas_dimention: Teknik

			MeasuresLevel
- Kode Akademik	Tahun Akademik	Media	Jumlah
All Th_akademik_dimenti	All Th_akademik_dimenti	All Promotion Dimention	632
		Internet	3
		Kampus Expo	9
		Keluarga	70
		Koran	1
		LAINYA	431
		Saudara	2
		Teman	116
+ 1011	1011 Total	All Promotion Dimention	205
		Internet	3
		Kampus Expo	2
		Keluarga	34
		Koran	1
		LAINYA	137
		Saudara	2
		Teman	26
+ 1112	1112 Total	All Promotion Dimention	214
		Internet	
		Kampus Expo	7
		Keluarga	36
		Koran	
		LAINYA	82
		Saudara	
		Teman	89
+ 1213	1213 Total	All Promotion Dimention	213
		Internet	

Gambar 2.9. OLAP Analisa Sebaran Mahasiswa Melalui Jalur Promosi.



Gambar 2.10. Grafik Sebaran Mahasiswa Melalui Jalur Promosi

Dari hasil OLAP pada gambar diatas diatas dapat dilihat dan diuraikan bahwa ada beberapa jumlah mahasiswa yang diterima di fakultas teknik Unitomo berdasarkan jalur promosi yang digunakan, diantaranya adalah sebagai berikut :

1) Total mahasiswa FT. adalah **632 mahasiswa**, total mahasiswa tersebut diperoleh selama tiga tahun akademik yang rata-rata dari masing-masing tahun akademik (TA)

mendapatkan 200-an mahasiswa. TA. **1011** yaitu **205 mahasiswa**, TA. **1112** sebanyak **214 mahasiswa**, dan TA. **1213** sebanyak **213 mahasiswa**.

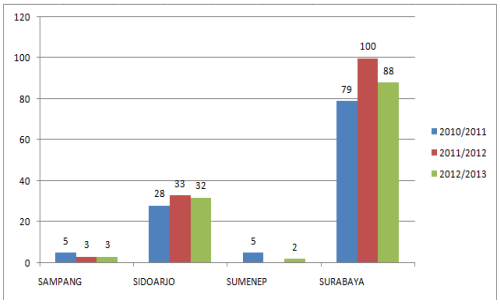
2) Dari **632 mahasiswa**, yang menduduki peringkat pertama yaitu melalui jalur **lainnya** sebanyak **431 mahasiswa** kemudian disusul **teman 116 mahasiswa**, **keluarga/saudara** sebanyak **72 mahasiswa**, **kampus expo 9 mahasiswa**, kemudian **internet** dan **Koran** masing-masing mendapatkan **3 mahasiswa** dan **1 mahasiswa**.

analisa data mahasiswa

Promotion Dimention	All Promotion Dimen	mahasiswa_dimention	All mahasiswa_dimer
status_dimention	All status_dimention	fakultas_dimention	Teknik

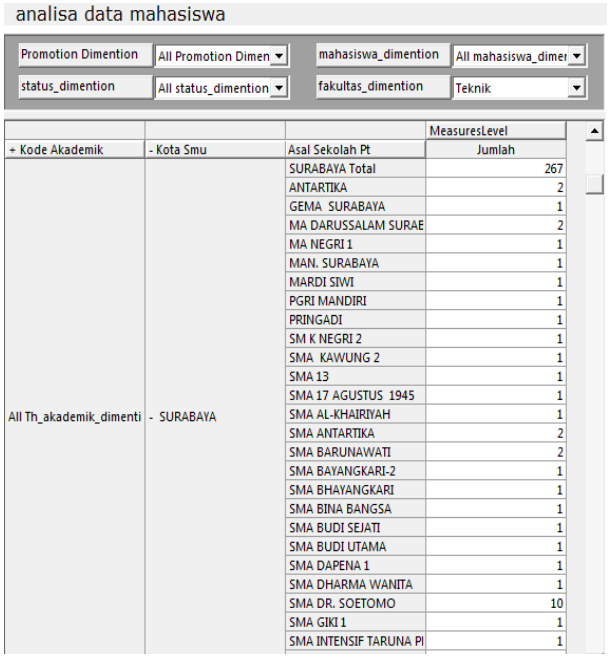
			MeasuresLevel
- Kode Akademik	Tahun Akademik	+ Kota Smu	Jumlah
All Th_akademik_dimenti	All Th_akademik_dimenti	+ PONTIANAK	1
		+ PROBOLINGGO	3
		+ PURBALINGGA	2
		+ REMBANG	1
		+ RUTENG	1
		+ SALATIGA	1
		+ SAMBAS	1
		+ SAMPANG	11
		+ SEMARANG	1
		+ SIDOARJO	93
		+ SITUBONDO	2
		+ SORONG	2
		+ SRAGEN	2
		+ SUMBA BARAT	1
		+ SUMBA TIMUR	1
		+ SUMBAWA	1
		+ SUMBAWA BESAR	1
		+ SUMENEP	7
		+ SURABAYA	267
		+ TARAKAN	1
		+ TIMOR TENG UTR	2
		+ TRENGGALEK	3
		+ TUBAN	8
		+ WAMENA	2
		+ WONOGIRI	1
		+ YOGYAKARTA	1

Gambar 2.11. OLAP Analisa Sebaran Mahasiswa Berdasarkan Wilayah



Gambar 2.12. Grafik Sebaran Mahasiswa Berdasarkan Wilayah

Dari hasil OLAP pada diatas dapat dilihat dan diuraikan bahwa ada beberapa jumlah mahasiswa yang diterima di fakultas teknik Unitomo di tahun yang sama berdasarkan wilayah (kabupaten-kota) yang masing-masing di dominasi wilayah **Surabaya** sebanyak **267 mahasiswa**, disusul **Sidoarjo** sebanyak **93 mahasiswa**, dan pulau **Madura** (sampang 11 dan sumenep 7) total sebanyak **18 mahasiswa**.



Gambar 2.13. OLAP Analisa Sebaran Mahasiswa Berdasarkan Asal Sekolah

Dari hasil OLAP pada gambar diatas dapat dilihat dan diuraikan bahwa ada beberapa jumlah mahasiswa yang diterima di fakultas teknik Unitomo di tahun yang sama berdasarkan Asal SMU di kota Surabaya secara garis besar yang tertinggi berasal dari SMA DR. SOETOMO sebanyak 10 mahasiswa, SMA Antartika sebanyak 4 mahasiswa, SMA Barunawati 2 mahasiswa, MA Darussalam 2 mahasiswa dan sisanya dari sekolah lain 249 mahasiswa.

Daftar Pustaka

Forsman, Sarah., 1997., *OLAP Council White Paper*, OLAP Council.

Inmon, Wiliam. 2002., *Building the Data Warehouse.*, 3rd Edition, New York : Wiley

McLeod, R., Jr. and G.P.Schell. 2007, *Management Information Systems,* 10th ed., Upper Saddle River, pearson Prentice Hall

Microsoft Corp.(2001), *Data Mining with Microsoft SQL Server 2000 Technical Reference*, Microsoft Press, Redmond, Washington

Nolan, Sean And Huguelet, 2000. "Microsoft SQL Server 7.0 Data Warehousing Training Kit". Microsoft Prees, USA.

Ponniah, Paulraj. 2001., *Data Warehouse Fundamental: A Comprehensive Guide for IT Proffesionals.* Singapore, John Wiley&Sons, Inc.

Singh, Harry, 1999. *Interactive Data Warehousing*, Practice-Hall, Inc.,

Turban, Efraim and Jaye E. Aronson.1998. *Decission and Support Systems and Intelligent System,* Fifth, Prentice Hall International.