



**PEMBUATAN DATA WAREHOUSE PEMPROV DKI JAKARTA MENGGUNAKAN
PENTAHO DATA INTEGRATION KETTLE**

SKRIPSI

SYAIFUL ANWAR

1110512087

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
2018**



**PEMBUATAN DATA WAREHOUSE PEMPROV DKI JAKARTA MENGGUNAKAN
PENTAHO DATA INTEGRATION KETTLE**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
Sarjana Komputer**

SYAIFUL ANWAR

1110512087

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
2018**

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Syaiful Anwar
NIM : 1110512087
Tanggal : 15 Agustus 2018

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 15 Agustus 2018



PERNYATAAN PERSETUJUAN PUBLIKASISKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini :

Nama : Syaiful Anwar
NIM : 1110512087
Fakultas : Ilmu Komputer
Program Studi : Sistem Informasi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Pembuatan Data Warehouse Pemprov DKI Jakarta Menggunakan Pentaho Data Integration Kettle

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengafih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 15 Agustus 2018

Yang menyatakan,



(Syaiful Anwar)

PENGESAHAN

Dengan ini dinyatakan bahwa Tugas Akhir berikut:

Nama : Syaiful Anwar

NIM : 1110512087

Program Studi : Sistem Informasi

Judul Tugas Akhir : Pembuatan Data Warehouse Pemprov DKI Jakarta
Menggunakan Pentaho Data Integration Kettle

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.



Kraugusteeliana, S.Kom., M.Kom., M.M

Penguji 1



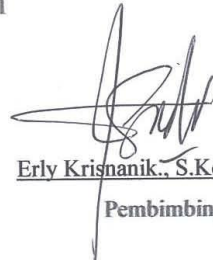
Anita Muliawati., S.Kom., M.T.I.

Penguji 2



Dr. Ermatita, M.Kom.

Dekan



Erly Krisnanik., S.Kom., M.M.

Pembimbing



Bambang Tri W., S.Kom., M.Si.

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 15 Agustus 2018

PEMBUATAN DATA WAREHOUSE PEMPROV DKI JAKARTA MENGGUNAKAN PENTAHO DATA INTEGRATION KETTLE

Syaiful Anwar

Abstrak

Data warehouse menyediakan cara yang efektif untuk melakukan analisis dan statistik terhadap data-data juga membantu dalam melakukan proses pengambilan keputusan. Skripsi ini mengenalkan *data warehouse* dan proses analisis *Online*. Menurut sistem informasi manajemen, skripsi ini menganalisis pemodelan *database* multidimensi yang berorientasi daya dan penyajian pandangan multidimensional. Sistem ini juga membangun gudang data berorientasi daya dengan menggunakan *Database Mysql*, *Software Open Source Pentaho Data Integration (Kettle)*, dan *Software BI tools Tableau* guna menampilkan sebuah *ouput* berupa *Dahboard*. Kemudian, Skripsi ini membahas data yang berasal dari gudang data Pemprov DKI Jakarta yang kemudian akan *diekstrak, transform, dan loading* (ETL). Pembahasan meliputi optimalisasi data *Open E-goverment* yang akan di buatkan *data warehouse* guna memperoleh hasil *Output* berupa *Dashboard* dengan menggunakan *Software BI Tools Tableau*. Alhasil sistem ini nantinya bisa membantu *Stakeholder* yang ada di lingkungan Pemprov DKI Jakarta untuk membantu dalam melakukan proses pengambilan keputusan.

Kata Kunci: *Data Warehouse, Power Management, OLAP, Stakeholders, System Analysis*

MAKING DATA WAREHOUSE PEMPROV DKI JAKARTA USING PENTAHO DATA INTEGRATION KETTLE

Syaiful Anwar

Abstract

The data warehouse provides an effective way to analyze and statistics on data, and helps in making decisions. This thesis introduces the data warehouse and online analysis process. According to the management information system, this thesis analyzes the modeling of power-oriented multidimensional databases and the presentation of multidimensional views. The system also builds power-oriented data warehouses using the Mysql database, Pentaho Open Source Data Integration (Kettle) Software, and Table Tools BI Tools to display an output in the form of a Dahboard. Then, this thesis discusses data coming from the DKI provincial government data warehouse which will then be extracted, transformed, and loaded (ETL). The discussion includes optimizing open government data that will be made in the data warehouse to obtain output in the form of a Dashboard using the Tableau BI Tools Software. As a result this system can help Stakeholders in the DKI Jakarta Provincial Government to assist the process of making decisions.

Keywords: Data Warehouse, Power Management, OLAP, Stakeholders, System Analysis

KATA PENGANTAR

Asslamu'allaikum Wr,Wb.

Puji dan syukur penulis panjatkan kehadiran Allah SWT, atas segala rahmat dan karunia-Nya sehingga laporan Skripsi ini berhasil diselesaikan dengan waktu yang telah ditentukan.

Pengerjaan laporan Skripsi ini dilakukan untuk memenuhi persyaratan untuk memperoleh gelar Sarjana pada Program Studi S1 Sistem Informasi, Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta.

Penulis ingin menyampaikan ucapan terima kasih kepada semua pihak yang telah memberikan dukungan dan bantuannya dalam penyelesaian laporan Skripsi ini. Adapun pihak-pihak tersebut diantaranya :

1. Kedua orangtua kami yang telah memberikan do'a dan dukungan tiada hentinya, dari sejak lahir hingga saat ini.
2. Ibu Dr. Ermatita, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta.
3. Ibu Erly Krisnanik S.Kom,. MM selaku dosen pembimbing saya yang telah membantu membimbing, mengarahkan sehingga Skripsi ini dapat selesai.
4. Bpk. Bambang Triwahyono, S.Kom., M.Si. selaku Ketua Program Studi S1 Sistem Informasi, yang telah banyak mendukung dan membantu mahasiswa/mahasiswi.
5. Bpk. Harry Susilo yang telah memberikan bantuan dalam pengumpulan data saat riset.
6. Bpk. Sayuti Bakrie, S.Kom,. M.Kom. selaku dosen PA yang selalu memberikan semangat kepada mahasiswa/i nya.
7. Teman-teman dan semua pihak yang telah membantu penulis dalam menyelesaikan Skripsi ini.

Jakarta, 15 Agustus 2018

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
PERNYATAAN ORISINALITAS	ii
PERNYATAAN PERSETUJUAN PUBLIKASI	iii
PENGESAHAN	iv
PERSETUJUAN SIDANG TUGAS AKHIR	v
ABSTRAK	vi
ABSTRACT.....	vii
KATAPENGANTAR	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR SIMBOL.....	xii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan dan Manfaat	3
1.3.1 Tujuan	3
1.3.2 Manfaat	4
1.4 Ruang Lingkup.....	4
1.5 Luaran Yang Diharapkan	4
1.6 Sistematika Penulisan.....	5
BAB 2 TINJAUAN PUSTAKA	6
2.1 Data, Informasi, dan Sistem Informasi	6
2.1.1 Data	6
2.1.2 Informasi	6
2.1.3 Sistem Informasi	6
2.2 Pengertian Basis Data	7
2.3 Sistem Basis Data.....	8
2.3.1 Database Management System	9
2.3.2 Database dan DBMS	9
2.3.3 Basis Data Relational	10
2.4 Konsep Dasar UML	12
2.4.1 Pengertian UML.....	12
2.4.2 Komponen UML	13
2.5 Data Warehouse	15
2.5.1 Karakteristik Data Warehouse	16
2.5.1.1 Berorientasi Subjek	16
2.5.1.2 Terintegrasi	17
2.5.1.3 Time Variant	18
2.5.1.4 Non Volatile.....	19
2.5.1.5 Berorientasi Proses.....	20
2.5.1.6 Mudah Digunakan.....	20
2.5.2 Manfaat Data Warehouse.....	21
2.6 Struktur Data Warehouse	21

2.7	Arsitektur Data Warehouse	22
2.7.1	Tujuan Arsitektur Data Warehouse.....	25
2.7.2	Delapan Layer pada Arsitektur Data Warehouse.....	25
2.7.3	Tiga Model Data Warehouse	27
2.8	Konsep Data Warehouse	27
2.9	Data Multidimensi Pada Data Warehouse	30
2.9.1	Data Mart	30
2.9.2	Hubungan Antara Data Mart dengan Data Warehouse.....	31
2.9.3	Perbedaan Antara Data Mart dan Data Warehouse.....	32
2.10	Data Multidimensi.....	33
2.10.1	Sepuluh Objek pada Data Multidimensi	33
2.10.2	Model Dimensional.....	35
2.10.3	Mengenal Multi Dimensional Database.....	38
2.11	OLAP dan OLTP.....	38
2.11.1	Empat Karakteristik OLAP	39
2.11.2	Empat Karakteristik OLTP	40
2.11.3	Perbedaan OLAP dan OLTP	41
2.12	ETL (Extract, Transform, Load)	41
2.13	Nine Step Design Method	42
2.14	Tools Pembuatan Data Warehouse	43
2.14.1	Database MySQL.....	43
2.14.2	Mengenal SQL (Structured Query Language)	44
2.14.3	Kelebihan MySQL	45
2.15	Bussiness Intelligence	45
2.15.1	Lima Konsep Dasar Bussiness Intelligence	47
2.15.2	Empat Level pada Bussiness Intelligence	48
2.15.3	Tableau Tools.....	49
2.16	Pentaho Data Integration.....	50
2.17	Analisi PIECES	50
BAB 3 METODOLOGI PENELITIAN.....		52
3.1	Kerangka Pikir	52
3.1.1	TahapanPengumpulan Data	52
3.1.2	Identifikasi Masalah.....	53
3.1.3	Perumusan Masalah	54
3.1.4	Analisa Kebutuhan.....	54
3.1.5	Perancangan	54
3.1.6	Uji Coba	54
3.1.7	Dokumentasi	54
3.2	Alat dan Bahan Yang Digunakan.....	54
3.3	Jadwal Penelitian.....	55
BAB 4 HASIL DAN PEMBAHASAN.....		56
4.1	Profil Pemprov DKI Jakarta.....	56
4.2	Visi dan Misi.....	57
4.2.1	Visi	57
4.2.2	Misi	57
4.3	Struktur Organisasi.....	58

4.3.1 Tugas dan Fungsi	58
4.4 Open Data Pemprov DKI Jakarta.....	60
4.5 Analisa Sistem Berjalan	60
4.5.1 Use Case Sistem Berjalan Kependudukan	62
4.5.2 Use Case Sistem Berjalan Dinas Pendidikan.....	62
4.6 Analisa Sistem Usulan	65
4.6.1 Metode Pengembangan Yang Digunakan.....	65
4.6.2 Flow Chart Sistem Usulan.....	67
4.6.3 Use Case Sistem Usulan.....	68
4.7 Class Diagram Sistem Usulan.....	69
4.7.1 Rancangan Struktur Data	70
4.7.2 Rancangan Kode	70
4.7.3 Proses Perencanaan	72
4.8 Perancangan Data Warehouse.....	72
4.8.1 Flow Chart Data Warehouse	72
4.8.2 Identifikasi Database Yang Digunakan.....	73
4.8.3 Arsitektur Data Warehouse	73
4.9 Perancangan Proses ETL.....	74
4.9.1 Proses ETL Dinas Kependudukan	75
4.9.2 Proses ETL Dinas Pendidikan.....	78
4.9.3 Proses Otomatisasi Job ETL	82
4.10 Analisa Query Report.....	83
4.10.1 Laporan Jumlah Siswa Putus Sekolah di Jakarta Selatan	84
4.10.2 Laporan Jumlah Penduduk Berdasarkan Jenis Kelamin	84
4.11 Halaman Interface Dashboard.....	85
4.11.1 Dashboard Jumlah Penduduk Sesuai Jenis Kelamin.....	85
4.11.2 Dashboard Jumlah Siswa Putus Sekolah.....	86
BAB 5 PENUTUP.....	87
5.1 Kesimpulan	87
5.2 Saran.....	87

DAFTAR PUSTAKA
RIWAYAT HIDUP
LAMPIRAN

DAFTAR TABEL





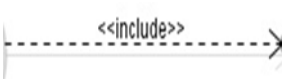

Tabel 3.3 Tabel Penelitian.....	55
Tabel 4.7.1 Tabel Struktur File Tabel	70
Tabel 4.8.2 Tabel Identifikasi Database Yang Digunakan	73
Tabel 4.9.1 Tabel Dinas Kependudukan	76
Tabel 4.9.2 Tabel Dinas Pendidikan	80

DAFTAR GAMBAR

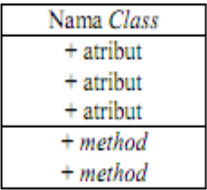




Gambar 2.5.1.1 Data Warehouse Berorientasi Subjek.....	16
Gambar 2.5.1.2 Data Warehouse Terintegrasi	17
Gambar 2.5.1.3 Data Warehouse Time Variant.....	18
Gambar 2.5.1.4 Data Warehouse NonVolatile.....	19
Gambar 2.6 Struktur Data Warehouse	21
Gambar 2.7 Arsitektur Data Warehouse	22
Gambar 3.1 Kerangka Pikir.....	52
Gambar 4.3 Struktur Organisasi.....	58
Gambar 4.5.1 Usecase Sistem Berjalan Dinas Kependudukan	62
Gambar 4.5.2 Usecase Sistem Berjalan Dinas Pendidikan	63
Gambar 4.5.3 Portal Opendatajakarta.go.id	63
Gambar 4.6.2 Flowchart Sistem Usulan	67
Gambar 4.6.3 Usecase Sistem Usulan.....	68
Gambar 4.7 Class Diagram Sistem Usulan	69
Gambar 4.8.1 Flowchart Data Warehouse	72
Gambar 4.8.3 Arsitektur Data Warehouse	73
Gambar 4.9 Perancangan Proses ETL.....	74
Gambar 4.9 Homepage Pentaho.....	74
Gambar 4.9.1 Database Staging Connection.....	75
Gambar 4.9.1 Proses ETL Dinas Kependudukan.....	76
Gambar 4.9.2 Proses ETL Dinas Pendidikan.....	79
Gambar 4.9.3 Proses Job ETL.....	83
Gambar 4.11.1 Dashboard Jumlah Penduduk Berdasarkan Jenis Kelamin	85
Gambar 4.11.2 Dashboard Jumlah Siswa Putus Sekolah.....	86

DAFTAR SIMBOL

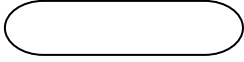
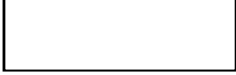
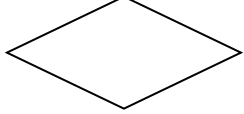

a. Simbol *Unified Modeling Language*
 1) Use Case Diagram

No.	Nama	Simbol	Penjelasan
1.	Actor		Actor atau pengguna sistem. Actor tidak terbatas hanya manusia saja, jika sebuah sistem berkomunikasi dengan aplikasi lain dan membutuhkan input atau memberikan output, maka aplikasi tersebut juga bisa dianggap sebagai actor.
2.	Use Case		Use case digambarkan sebagai lingkaran elips dengan nama use case dituliskan didalam elips tersebut.
3.	Association		Asosiasi digunakan untuk menghubungkan actor dengan use case. Asosiasi digambarkan dengan sebuah garis yang menghubungkan antara Actor dengan Use Case.
4.	Association		Asosiasi antara aktor dan use case yang menggunakan panah terbuka untuk mengindikasikan bila aktor berinteraksi secara pasif dengan sistem.
5.	Include		Include merupakan di dalam use case lain atau pemanggilan use case oleh use case lain.
6.	Extend		Extend merupakan perluasan dari use case lain jika kondisi atau syarat terpenuhi.

2) Class Diagram

No	Nama	Simbol	Keterangan
1.	Class		Class adalah blok-blok pembangun pada pemrograman berorientasi objek. Sebuah class digambarkan sebagai sebuah kotak yang terdiri atas 3 bagian, bagian tengah mendefinisikan property/atribut class. Bagian akhir mendefinisikan method-method dari sebuah class.
2.	Association	<p>1..* Owned by 1..1</p> 	Sebuah Asosiasi merupakan sebuah relationship paling umum antara 2 class dan dilambangkan oleh sebuah garis yang menghubungkan antara 2 class garis ini bisa melambangkan tipe-tipe relationship dan juga dapat menampilkan hukum-hukum multiplisitas pada sebuah relationship. (Contoh : One-to-one, one-to-many, many-to-many)
3.	Composition		Jika sebuah class tidak bisa berdiri sendiri dan harus merupakan bagian dari class yang lain, maka class tersebut memiliki relasi composition terhadap class tempat dia bergantung tersebut. Sebuah relationship composition digambarkan sebagai garis dengan ujung berbentuk jajaran genjang berisi/solid.
4.	Dependency		Kadang kala sebuah class diagram menggunakan class yang lain. Hal ini disebut dependency. Umumnya penggunaan dependency digunakan untuk menunjukkan operasi pada suatu class yang menggunakan class yang lain. Sebuah dependency dilambangkan sebagai sebuah panah bertitik-titik.
5.	Aggregation		Aggregation mengidikasikan keseluruhan bagian relationship dan biasanya disebut sebagai relasi.

b. Simbol *Flowchart*

No	Nama	Simbol	Keterangan
1.	Terminator		Simbol terminator (Mulai/selesai) merupakan tanda bahwa sistem akan dijalankan atau berakhir
2.	Proses		Simbol yang digunakan untuk melakukan pemrosesan data baik oleh user maupun komputer (sistem)
3.	Verifikasi		Simbol yang digunakan untuk memutuskan apakah valid atau tidak validnya suatu kejadian.
4.	Data		Simbol yang digunakan untuk mendeskripsikan data yang digunakan. Laporan : Simbol yang digunakan untuk menggambarkan laporan.