

PROPOSAL PROYEK

12S3202 – GUDANG DATA DAN KECERDASAN BISNIS

Perancangan *Data mart* Pengelolaan Keuangan Daerah

(Studi Kasus Pemerintahan Provinsi XYZ)



Disusun oleh:

1. 12S17027 – Stella Sitinjak
2. 12S17029 – Silvany Lumban Gaol
3. 12S17050 – Kotrel Manurung
4. 12S17056 – Maria Manullang

FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO

INSTITUT TEKNOLOGI DEL

APRIL 2020

DAFTAR ISI

DAFTAR ISI	i
1. Pendahuluan	1
1.1. Latar Belakang	1
1.2. Tujuan	2
1.3. Ruang Lingkup	2
2. Landasan Teori	3
2.1. <i>Data warehouse</i>	3
2.2. Data Mart	4
2.3. Arsitektur <i>Data warehouse</i>	5
2.4. Infrastruktur <i>Data warehouse</i>	6
2.3. Teknik Pemodelan Dimensional Data Warehouse	6
2.3.1. Proses Bisnis	6
2.3.2. Grain	7
2.3.3. Tabel Dimensi	7
2.3.4. Tabel Fakta	7
2.4. Star Schema	7
2.4. <i>Extract Transform and Load (ETL)</i>	8
2.5. <i>Online Analytical Processing (OLAP)</i>	8
3. Referensi	9

1. Pendahuluan

Pada bab ini akan dijelaskan mengenai latar belakang, tujuan, dan ruang lingkup dari proyek Perancangan *Data Mart* Pengelolaan Keuangan di Provinsi XYZ yang akan dikerjakan lebih lanjut agar menjadi acuan dalam pengerjaan proyek.

1.1. Latar Belakang

Terbitnya UU. No. 33 tahun 2004 tentang perimbangan keuangan antara pusat dan daerah, merupakan tanda bahwa reformasi keuangan daerah telah dimulai dengan kewenangan daerah yang semakin tinggi untuk mengurus dan mempertanggungjawabkan kinerja keuangannya secara mandiri, nyata, optimal, terpadu dan dinamis.

Banyak metode analisis yang dapat digunakan untuk menggambarkan kinerja keuangan daerah dan bermanfaat untuk melakukan pengukuran kinerja keuangan. Beberapa analisis yang paling umum digunakan adalah analisis rasio, trend dan peramalan (*forecasting*). Kebutuhan akan analisis untuk menyediakan informasi yang berkualitas dewasa ini terus meningkat. Selain permintaan terhadap data historis, data yang bersih, terkonsolidasi, dan sesuai waktu, permintaan para pengguna terhadap akses ke data *real-time*, tak terstruktur, dan atau data *remote*, juga kian meningkat. Sayangnya berbagai data yang dibutuhkan dalam proses analisis keuangan daerah tidak diorganisasikan dengan baik.

Dalam prakteknya, banyak ditemui data tersedia dalam *file* dan dokumen terpisah, diletakkan di gudang dan banyak yang kemudian hilang atau rusak. Data yang tidak diorganisasikan dengan baik menyebabkan kualitas *output* informasi yang dihasilkan dalam bidang keuangan tidak komprehensif sehingga hanya mencukupi kewajiban administratif, tanpa makna di tahun-tahun berikutnya. Padahal data yang terkumpul dari waktu ke waktu secara historis sangat dibutuhkan dalam membaca trend dan analisis rasio keberhasilan kinerja keuangan daerah saat ini dan untuk perencanaan tahun mendatang.

Kami melihat bahwa berbagai data yang dibutuhkan dalam proses analisis keuangan daerah tidak diorganisasikan dengan baik. Dalam prakteknya, banyak ditemui data tersedia dalam *file* dan dokumen terpisah, diletakkan di gudang dan banyak yang kemudian hilang atau rusak. Data yang tidak diorganisasikan dengan baik ini menyebabkan kualitas *output* informasi yang dihasilkan dalam bidang keuangan tidak komprehensif sehingga hanya mencukupi kewajiban administratif, tanpa makna di tahun-tahun berikutnya.

Kami mengusulkan desain sistem *data warehouse* dalam pengelolaan sistem keuangan daerah dan prototipe agar dapat membantu pemerintah daerah melakukan pengumpulan (*collecting*) data, mengorganisasi data dan menyediakan alat (*tool*) analisis yang dapat divisualisasikan dalam bentuk grafik, *dashboard* dan laporan informasi yang bermanfaat bagi manajemen dalam pengambilan keputusan terutama tentang keberlangsungan pembangunan daerah.

1.2. Tujuan

Tujuan dari pengerjaan proyek ini adalah sebagai berikut:

1. Merancang data *warehouse* untuk mempermudah daerah provinsi XYZ dalam menyimpan data setiap pemasukan dan pengeluaran keuangan daerah di provinsi XYZ.
2. Untuk dapat mengontrol jumlah pemasukan dan pengeluaran pada provinsi XYZ.
3. Untuk dapat mengorganisasikan berbagai data yang dibutuhkan dalam proses analisis keuangan daerah.
4. Untuk dapat membantu pemerintah daerah melakukan pengumpulan (*collecting*) data.
5. Untuk dapat membantu pemerintah daerah dalam menyediakan alat (*tools*) analisis yang dapat divisualisasikan (dalam bentuk grafik, *dashboard*, dan laporan informasi).

1.3. Ruang Lingkup

Ruang lingkup proyek ini adalah Pemerintahan provinsi terutama yang memiliki kinerja keuangan (pemasukan dan pengeluaran) yang tidak terorganisir dengan baik. Proyek ini dapat membantu pemerintah daerah dalam melakukan pengumpulan (*collecting*) data, mengorganisasi data dan menyediakan alat (*tools*) analisis yang dapat divisualisasikan dalam bentuk grafik, *dashboard* dan laporan informasi.

2. Landasan Teori

Pada bab ini akan dijelaskan mengenai teori-teori yang digunakan sebagai landasan dalam pengerjaan proyek.

2.1. Data warehouse

Data warehouse merupakan suatu paradigma baru di lingkungan pengambilan keputusan strategis. *Data warehouse* bukan suatu produk tetapi suatu lingkungan dimana pengguna dapat menemukan informasi strategis [Poniah, 2001, h.14]. *Data warehouse* merupakan kumpulan data-data logik yang terpisah menggunakan *database* operasional dan merupakan suatu ringkasan. Adapun karakteristik dari *data warehouse* [Poniah, 2001,h.20-24] adalah sebagai berikut.

1. Berorientasi subyek

Data warehouse merupakan tempat penyimpanan berdasarkan subyek bukan berdasarkan aplikasi. Subyek merupakan bagian dari suatu perusahaan. Contoh subyek dalam perusahaan manufaktur adalah penjualan, konsumen, inventori, dan lain sebagainya. Tabel dibawah ini merupakan perbedaan mengenai *data warehouse* dan *database* operasional.

Tabel 1 Perbedaan *Database* Operasional dan *Data Warehouse*

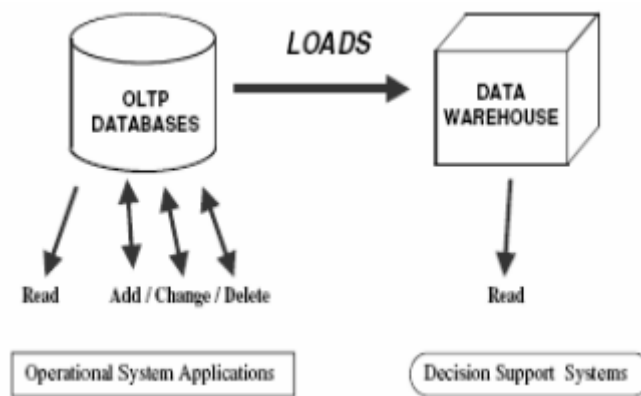
	<i>Database</i> Operasional	<i>Data warehouse</i>
Isi data	Terbaru (<i>up-to-date</i>)	Arsip, <i>history</i> , rangkuman
Struktur Data	Dioptimasi untuk transaksi dan normalisasi	Dioptimasi untuk <i>query</i> yang kompleks dan sangat jarang di normalisasi
Frekuensi Akses	Tinggi	Sedang-rendah
Tipe Akses	<i>Read, update, delete</i>	<i>Read</i>
Penggunaan	<i>Update</i> secara terus menerus	<i>Update</i> secara periodic
<i>Users</i>	Banyak	Lebih sedikit

2. Data yang terintegrasi

Sumber data yang ada dalam *data warehouse* tidak hanya berasal dari *database* operasional (*internal source*) tetapi juga berasal dari data diluar sistem (*external source*). Data pada sumber berbeda dapat di-*encode* dengan cara yang berbeda. Sebagai contoh, data jenis kelamin dapat diencode menjadi 0 dan 1 pada satu tempat dan “m” dan “f” pada tempat lain.

3. Non-volatile

Data dalam *database* operasional akan secara berkala atau periodik dipindahkan kedalam *data warehouse* menggunakan jadwal yang sudah ditentukan. Misalnya perhari, perminggu, perbulan, dan lain sebagainya. Sekali masuk ke dalam *data warehouse*, data berstatus *read-only*. Pada gambar 2.1 dibawah ini bisa dilihat bahwa *database* OLTP bisa dibaca, di-*update*, dan dihapus. Tetapi pada *database data warehouse* hanya bisa dibaca.



Gambar 1 Data warehouse bersifat non-volatile

4. Time-Variant

Sistem operasional mengandung data yang bernilai sekarang sedangkan data dalam *data warehouse* mengandung data tidak hanya data terkini tetapi juga data *history* yang akan digunakan dalam analisis dan pengambilan keputusan. Waktu adalah dimensi penting yang harus didukung oleh semua *data warehouse*. Data untuk analisis dari berbagai sumber berisi berbagai nilai waktu, misalkan harian, mingguan, dan bulanan.

5. Ringkas

Jika diperlukan, data operasional dikumpulkan ke dalam ringkasan-ringkasan.

6. Granularity

Pada sistem operasional data dibuat secara *real-time* sehingga untuk mendapatkan informasi langsung dilakukan proses *query*. Pada *data warehouse* menganalisis harus memperhatikan *level-of-detail* misalkan perhari, ringkasan perbulan, ringkasan per-tiga-bulan.

7. Tidak ternormalisasi

Data di dalam sebuah *data warehouse* biasanya tidak ternormalisasi dan sangat redundan.

2.2. Data Mart

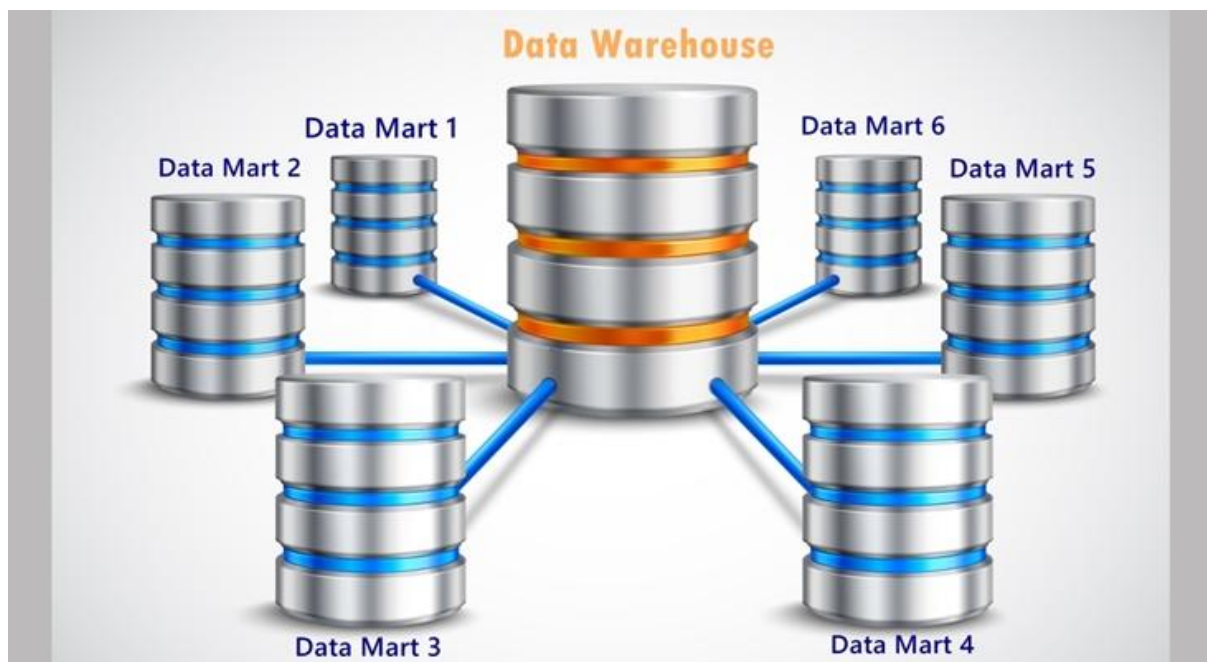
Menurut Wang (2006) “*data mart* adalah sebuah *data warehouse* yang terbatas pada lingkup dan fasilitas dan untuk sebuah domain tertentu”. Dapat dikatakan bahwa *data mart* adalah sebuah *subset* dari keseluruhan *data warehouse* suatu perusahaan ataupun organisasi. Berikut perbedaan karakteristik antara *data warehouse* dengan *data mart*.

Tabel 2. Perbandingan Karakteristik Data Warehouse dengan Data Mart

Data Warehouse	Data Mart
Subyek penggunaannya luas (Multiple)	Subyek penggunaannya departemental (Single)
Merupakan kumpulan <i>data mart</i>	Merupakan proses bisnis tunggal
Waktu implementasi berbulan-bulan bahkan hingga bertahun-tahun.	Waktu implementasi berbulan-bulan

Dibuat untuk menampilkan data perusahaan	Dibuat untuk menampilkan data per departemen.
Menampung seluruh data perusahaan	Menampung data per departemen.
Ukuran data 100 GB sampai > 1 TB	Ukuran data <100 GB

Tronthwaite (2011), menyatakan definisi *data mart* sedikit lebih teknis, yaitu merupakan sebuah *subset* dari *data warehouse* yang biasanya berkaitan dengan sebuah fungsi bisnis, dan mungkin hanya konsentrasi pada beberapa tabel fakta, tabel dimensi untuk kebutuhan *query* dan laporan.

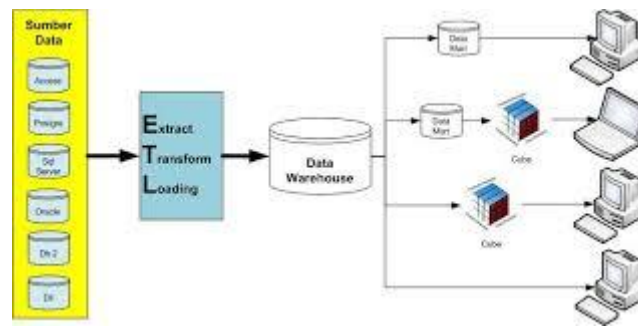


Gambar 2 Perbedaan *data warehouse* dengan *data mart*

Dari gambar, dapat kita lihat bahwa, data mart merupakan *subset* dari sebuah *data warehouse* yang dimiliki suatu organisasi. Perubahan yang terjadi pada *data mart* tidak akan mempengaruhi *data warehouse* ataupun *data mart* lainnya.

2.3. Arsitektur *Data warehouse*

Arsitektur data warehouse adalah sekumpulan aturan atau struktur yang menyediakan sebuah framework untuk keseluruhan desain dari sistem tertentu. Sebuah arsitektur data menyediakan sebuah framework yang dapat mengidentifikasi dan mengerti bagaimana data akan bergerak melalui sistem dan keperluan dalam perusahaan.



Gambar 3 Arsitektur Data Warehouse

Gambar di atas mendeskripsikan aliran data dari sumber data operasional ke *data warehouse* melalui proses transformasi. Gambar di atas juga menunjukkan posisi *data warehouse* yang terdiri dari sumber data, ETL, *data mart* dan *cube* kemudian dihantarkan kepada pengguna. Sumber data merupakan data operasional yang disimpan dalam *database*, yang akan diproses oleh ETL dan diintegrasikan ke dalam *data warehouse*. Sedangkan *data mart* dan *cube* berisi data yang mendukung fungsi bisnis (Inmon, 2005). *Data mart* adalah sub-set dari *data warehouse* yang umumnya terdiri dari sebuah subjek tunggal. Dapat diartikan juga sebagai serangkaian data yang hanya menjelaskan satu fungsi dari operasi perusahaan.

2.4. Infrastruktur Data warehouse

Infrastruktur data *warehouse* terdiri dari *software*, *hardware*, pelatihan-pelatihan dan komponen-komponen lainnya yang memberi dukungan yang diperlukan dalam mengimplementasikan arsitektur data *warehouse* (Poe, 1998). Salah satu instrument yang mempengaruhi keberhasilan pengembangan data *warehouse* adalah pengidentifikasian arsitektur mana yang terbaik dan infrastruktur yang dibutuhkan. Arsitektur dan infrastruktur sangat erat hubungannya. Arsitektur yang sama mungkin memerlukan infrastruktur yang berbeda tergantung pada lingkungan perusahaan ataupun organisasi.

2.3. Teknik Pemodelan Dimensional Data Warehouse

Menurut (Kimball & Ross, The Data Warehouse Toolkit, 2000) terdapat 4 langkah untuk melakukan pemodelan dimensional, yaitu:

1. Pemilihan proses bisnis
2. Deklarasi Grain
3. Identifikasi Dimensi
4. Identifikasi Fakta

2.3.1. Proses Bisnis

Proses bisnis merupakan aktivitas operasional yang dilakukan dalam suatu organisasi seperti: memproses pendistribusian gaji pegawai, menerima pesanan, dan lainnya. Proses bisnis menghasilkan matriks kinerja yang akan menerjemahkan fakta kedalam tabel fakta. Pemilihan proses bisnis sangat penting karena nantinya akan berpengaruh pada target desain data warehouse, grain, dimensi serta fakta yang digunakan.

2.3.2. Grain

Granularity (Grain) merupakan tingkat detail data dalam suatu data *warehouse*. Semakin detail suatu data, maka akan semakin rendah tingkat *granularity* nya. jadi, data yang semakin detail (tingkat kedetailan tinggi) akan memiliki *Level* yang paling rendah, misalnya pada data transaksi.

Grain penting dalam *data warehouse* , karena:

1. Semakin rendah level *granularity*, maka tingkat kedetailan data akan semakin besar dan berpengaruh terhadap pertanyaan yang akan dijawab.
2. Semakin rendah level *granularity*, maka jumlah data yang disimpan dalam *data warehouse* juga akan semakin besar.

Jika di analisa dalam penggunaannya, *data granularity* bertepatan pada level yang berbeda bergantung pada tingkat keperluan organisasi tersebut.

2.3.3. Tabel Dimensi

Tabel dimensi (*dimension table*) berisi atribut penjelasan yang digunakan pada *business intelligence* untuk mengelompokkan dan mem-filter fakta. Tabel dimensi mengandung konteks “apa, siapa, kapan, dimana dan bagaimana” yang terjadi pada aktivitas bisnis. Dengan mendefinisikan *granularity* pada tabel fakta, maka dimensi apa saja yang diperlukan dapat teridentifikasi juga. Tabel dimensi juga diistilahkan sebagai “nyawa” dari *business intelligence*, karena menjadi alat utama dalam menjelaskan konteks dalam analisis fakta di peristiwa bisnis (Kimball, Kimball Dimensional Modelling Techniques, 2013)

2.3.4. Tabel Fakta

Tabel Fakta (*Fact table*) merupakan tabel yang mengandung sesuatu yang dapat diukur seperti jumlah, harga, dan lainnya. Tabel fakta juga merupakan kumpulan dari *primary key* dan *foreign key* yang dimiliki oleh masing-masing tabel dimensi. tabel fakta juga mengandung data yang bersifat historis (Saptadi & Lisangan, 2010)

2.4. Star Schema

Pada sistem OLTP (*Online Transactional Processing*) digunakan suatu teknik pemodelan data yang disebut sebagai E-R (*Entity-Relationship*). Pada *data warehouse* digunakan teknik pemodelan data yang disebut *dimensional modelling technique*. Pemodelan dimensional adalah suatu model berbasis pemanggilan yang mendukung akses *query* volume tinggi. *Star Schema* adalah alat dimana pemodelan dimensional diterapkan dan berisi sebuah tabel fakta pusat.

Tabel fakta berisi atribut deskriptif yang digunakan untuk proses *query* dan *foreign key* untuk menghubungkan ke tabel dimensi. Atribut analisis keputusan terdiri dari ukuran performa, metrik operasional, ukuran agregat, dan semua metrik yang lain yang diperlukan untuk menganalisis performa organisasi. Tabel fakta menunjukkan apa yang didukung oleh *data warehouse* untuk analisis keputusan. Tabel dimensi mengelilingi tabel fakta pusat. Tabel dimensi berisi atribut yang menguraikan data yang dimasukkan dalam tabel fakta. Tabel dimensi menunjuk bagaimana data akan dianalisis.

Pada skema ini, hubungan antara tabel dimensi dan tabel fakta menyerupai bintang, di mana satu tabel fakta dihubungkan dengan beberapa tabel dimensi. Query yang terbentuk antara tabel fakta dan sejumlah tabel dimensi dinamakan star query. Setiap tabel dimensi direlasikan secara langsung dengan tabel fakta berdasarkan kolom primary key dan foreign key, namun diantara tabel dimensi tidak ada yang saling berelasi hal ini membuat proses eksekusi query akan lebih optimal.

2.4. Extract Transform and Load (ETL)

ETL merupakan proses pemasukan data operasional kedalam *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 proses ETL adalah data yang memenuhi kriteria *data warehouse*, seperti data historis, terpadu, terangkum, statis, dan memiliki struktur yang dirancang untuk keperluan proses analisis.

1. Extract

Langkah pertama pada proses ETL adalah mengekstrak data dari sumber-sumber data. Kebanyakan proyek *data warehouse* menggabungkan data dari sumber-sumber yang berbeda. Sistem-sistem yang terpisah sangat mungkin menggunakan format data yang berbeda. Ekstraksi bertujuan untuk mengubah data dengan format yang ada ke dalam suatu format yang berguna untuk proses transformasi.

2. Transform

Tahap transformasi menggunakan serangkaian aturan atau fungsi untuk merubah data sumber agar sesuai dengan struktur data tujuan yakni *data warehouse*.

3. Load

Tahap *load* berfungsi untuk memasukkan data ke dalam suatu *data warehouse*. Jangka waktu proses ini tergantung pada kebutuhan organisasi. Pada beberapa *data warehouse*, proses *load* dapat dilakukan setiap minggu untuk mengisi keseluruhan informasi yang ada secara kumulatif, sementara pada *data warehouse* yang lain, proses *load* dilakukan dengan menambahkan data baru dalam suatu bentuk yang *historical* (setiap jam). Waktu dan jangkauan penggantian atau penambahan data tergantung pada perancangan *data warehouse* pada saat menganalisis keperluan informasi.

2.5. Online Analytical Processing (OLAP)

Online Analytical Processing (OLAP) merupakan salah satu tools yang dapat digunakan dalam mengakses informasi dalam *data warehouse* secara efektif untuk proses inline analysis, OLAP dapat memberikan respon yang cepat berupa jawaban terhadap kueri analitik yang kompleks. Multidimensional data model dan teknik agregasi data yang dimiliki oleh OLAP dapat mengatur dan membuat kesimpulan dari data dalam jumlah besar, sehingga dapat dievaluasi secara cepat dengan menggunakan online analysis dan graphical tool. Sistem OLAP menyediakan kecepatan dan fleksibilitas untuk melakukan support analisis secara real time.

3. Referensi

- [1] Andreas Ronald, D.S. (2010). Analisis Kinerja Keuangan dan Pertumbuhan Ekonomi sebelum dan sesudah diberlakukannya Otonomi Daerah di Kabupaten Kulon Progo. *Bisnis dan Ekonomi* Vol. 1 No. 1, 31-42.
- [2] Halim, A. (2001). *Bunga Rampai : Manajemen Keuangan Daerah* (Pertama ed.): UPP AMP YKPN.
- [3] H. Jusuf dan A. Azimah, “Perancangan *Data warehouse* pada Perpustakaan Universitas Nasional,” *Jurnal Basis Data, ICT Research Center UNAS*, vol 3, no. 1, pp. 2-3, 2008.
- [4] K. Haryono, “Academia,” [Online]. Available:
https://www.academia.edu/7312328/Jurnal_Warehousing_Kholid_Haryono?auto=download. [Diakses pada 13 April 2020].
- [5] Mahsu, M. (2009). *Pengukuran Kinerja Sektor Publik*, Yogyakarta: BPFE UGM Yogyakarta.
- [6] Thesaurianto, K. (2007). *Analisis Pengelolaan Keuangan Daerah terhadap Kemandirian Daerah*, Semarang: Program Pasca Sarjana Universitas Diponegoro.
- [7] V. Poe, *Building Data Warehouse for Decision Support*, Prentice Hall: edisi-2, 1998.