

## ANALISA DAN PERANCANGAN DATA WAREHOUSE DENGAN METODE NINE STEP KIMBALL DI PT SURGANYA MOTOR INDONESIA

*Dedy Prastyo<sup>1</sup>, Aji Supriyanto<sup>2</sup>*

<sup>1,2</sup>Program Studi Teknik Informatika, Fakultas Ilmu Teknologi Informasi, Universitas Stikubank Semarang  
e-mail: <sup>1</sup>me@dedyprastyo.com, <sup>2</sup>ajisup@edu.unisbank.ac.id

**ABSTRAK**

Informasi menjadi salah satu asset yang berharga untuk perusahaan. Bisnis memerlukan pengambilan keputusan yang berorientasi pada data yang valid, akurat dan cepat Informasi yang akurat jika digunakan dengan tepat dapat menjadi keunggulan kompetitif bagi perusahaan. Teknologi data warehouse dapat membantu perusahaan dalam menangani data dalam jumlah besar yang berguna bagi pemangku kepentingan dalam proses pengambilan keputusan. Penelitian ini bertujuan untuk memperlihatkan potensi data warehouse pada perusahaan sehingga perusahaan mendapat nilai lebih dari informasi yang dimiliki. Penelitian ini menggunakan data warehouse dan nine step Kimball untuk pengumpulan data terkait penjualan dan Tableau untuk melakukan visualisasi data sehingga pemangku kepentingan dapat melihat posisi penjualan secara cepat dan dapat mengambil keputusan terkait kondisi terkini dengan tepat. Hasil penelitian ini adalah sistem data warehouse yang dapat melakukan proses analisa data penjualan dan sebuah dashboard yang dapat digunakan untuk memantau kondisi penjualan. Hasil pengujian performa pada query MDX terhadap cube menunjukkan rata-rata waktu yang dibutuhkan untuk eksekusi query adalah 1 detik.

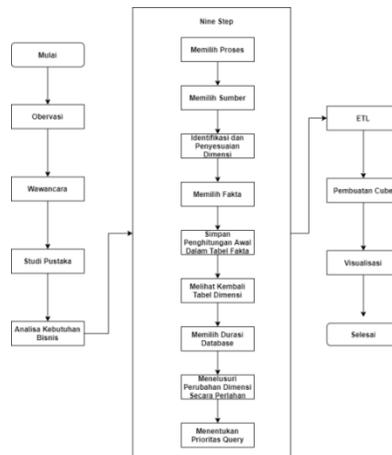
**Kata Kunci:** data warehouse, OLAP, multidimensional data, kimball, nine-step, tableau

**1. PENDAHULUAN**

Bisnis memerlukan pengambilan keputusan yang dilakukan dengan cepat dan tepat. Kunci dalam sukses atau tidaknya pengambilan keputusan bergantung pada kualitas dan kecepatan dalam pengambilan keputusan [1]. Informasi merupakan hal terpenting dalam perusahaan [2]. Pengambilan keputusan berdasarkan data (*data-driven decision making*) berkaitan erat dengan peningkatan produktivitas [3]. Tantangan utama bagi bisnis adalah untuk mendapatkan informasi yang tepat disaat yang tepat oleh orang yang tepat dengan cara semudah mungkin. [4].

**2. TINJAUAN PUSTAKA**

Data merupakan elemen dari informasi yang sudah dibentuk ke dalam sebuah bentuk yang bermanfaat dan dapat digunakan untuk manusia [5]. Data dan fakta yang terorganisir memiliki nilai tambah selain nilai fakta individu [6]. OLAP (*Online Analytical Processing*) merupakan merupakan salah satu teknik dalam analisa data yang dapat meningkatkan kualitas dalam proses pengambilan keputusan [7]. Data dalam OLAP *cube* disimpan dalam bentuk multidimensi. Data *warehouse* memungkinkan untuk memperoleh informasi yang diperlukan dalam melakukan analisis lebih jauh terhadap subjek-subjek tertentu secara lebih cepat [8]. Data yang telah didapat divisualisasikan menggunakan komputer untuk menemukan metode terbaik dalam menampilkan data untuk dengan cara penerimaan alami manusia serta memberikan cara untuk melihat data yang sulit dilihat dengan pemikiran [9].



Gambar 1. Metode Penelitian

### 3. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini meliputi observasi, wawancara, studi pustaka, proses implementasi nine step Kimball, proses pembuatan cube dan visualisasi data.

#### 3.5. 3.1. Observasi

Proses dimulai dengan melakukan observasi. Observasi dilakukan di toko dan kantor cabang. Hal ini untuk mengetahui kegiatan yang dilakukan terkait proses transaksi.

#### 3.6. 3.2. Wawancara

Proses dilanjutkan dengan wawancara dengan stakeholder perusahaan. Tujuan dari proses ini untuk mendapatkan sudut pandang dari pemangku kepentingan di perusahaan. Wawancara juga bertujuan untuk menguatkan hasil yang didapat selama proses observasi.

#### 3.7. 3.3. Studi Pustaka

Studi Pustaka dilakukan untuk mempelajari hal yang berkaitan dengan data warehouse, metode nine step dari Kimball, OLAP (Online Analytical Processing) dan visualisasi data.

#### 3.8. 3.4. Analisa Kebutuhan Bisnis

Pada tahap ini dilakukan analisa lebih lanjut proses bisnis di perusahaan terkait dengan penggunaan data penjualan. Analisa meliputi proses penjualan, proses retur barang, proses penerimaan barang yang terjadi di toko dan *warehouse*. Tujuan agar apa yang dianalisa sesuai dengan kebutuhan *stakeholder* perusahaan.

#### 3.9. 3.5. Proses Nine Step Kimball

Metode pembuatan data warehouse yang digunakan dalam penelitian ini adalah metode nine step dari Kimball. Metode ini memiliki sembilan tahapan yaitu:

##### *Choose the Process*

Tahapan pertama adalah menentukan subjek dari masalah yang dihadapi. Tentukan proses yang penting dan ingin diolah dari kegiatan operasional. Proses yang akan diteliti adalah proses penjualan yang terjadi.

##### *Choose the Grain*

*Grain* merupakan data dari tabel fakta yang dapat dianalisis. Langkah selanjutnya adalah memilih data yang dapat dianalisis.

##### *Identify and Confirm the Dimension*

Mengidentifikasi dimensi secara detail untuk mendeskripsikan sesuatu.

##### *Choose the Fact*

Memilih fakta-fakta yang akan digunakan. Fakta-fakta tersebut memiliki data yang dapat dihitung dan direpresentasikan kedalam bentuk laporan maupun visualisasi.

##### *Store Pre-calculation In the Fact Table*

Setelah fakta-fakta tersebut dipilih kemudian dikaji ulang untuk menentukan peluang untuk digunakan dalam perhitungan.

##### *Rounding Out the Dimension Table*

Mengembalikan fakta yang telah dihitung ke dalam tabel dimensi.

##### *Decide the Duration of Database and Periodicity of Update*

Menentukan batas waktu dari umur data yang diambil dan akan dipindahkan ke tabel fakta. Penentuan batas ini dengan memperhatikan tingkat akurasi yang dimiliki oleh data historis.

##### *Track Slowly Changing Dimension*

Lacak perubahan pada tabel dimensi. Perubahan pada tabel dimensi dapat dilakukan dengan tiga cara yaitu mengganti secara langsung pada tabel dimensi, membentuk *record* baru pada setiap perubahan data dan membentuk kolom baru untuk setiap perubahan data.

##### *Decide the Query Priorities and Query Models*

Langkah terakhir adalah menentukan prioritas dan model *query* pada proses ETL (*extract, transform, load*). *Query* menentukan proses mana saja yang dijalankan lebih dahulu selama proses ETL.

#### 3.10. 3.6. ETL (Extract, Transform, Load)

Proses ETL dilakukan untuk memindahkan data dari database operasional ke dalam database data warehouse. Proses ETL dilakukan secara berkala guna memastikan data yang tersimpan di data warehouse merupakan data terbaru.

#### 3.11. 3.7. Pembuatan Cube

Proses OLAP dimulai dengan pengambilan data dari database warehouse. Data dari database warehouse kemudian diubah menjadi dimensi-dimensi yang dikumpulkan sedemikian rupa sehingga dapat diakses dan dianalisa. Hasil akhir dari proses OLAP ini adalah terbentuknya sebuah cube yang dapat dianalisa menggunakan metode seperti summarization (peringkasan), consolidation (konsolidasi) dan aggregation (agregasi).

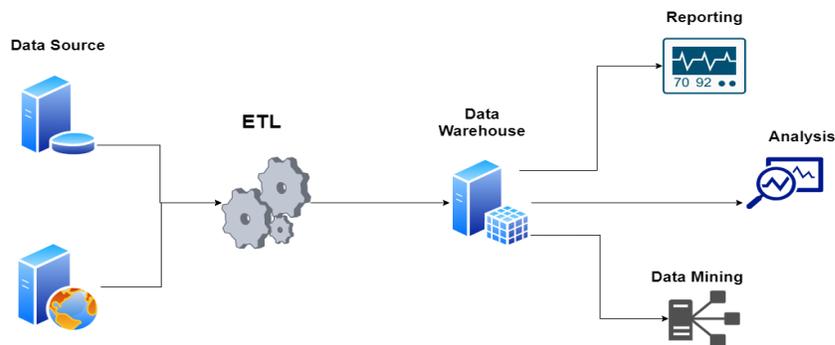
#### 3.8 Visualisasi Data

Proses visualisasi data dilakukan dengan mengambil data dari cube dalam database kemudian data yang berwujud angka diubah ke dalam bentuk visual agar mudah dipahami oleh pemangku kepentingan (stakeholder) untuk memudahkan proses pengambilan keputusan berdasarkan data (data driven decision). Proses visualisasi data menggunakan tools untuk memudahkan pengguna dalam melihat hasil visualisasi data tanpa harus memiliki pengetahuan teknis tentang database.

**4. HASIL DAN PEMBAHASAN**

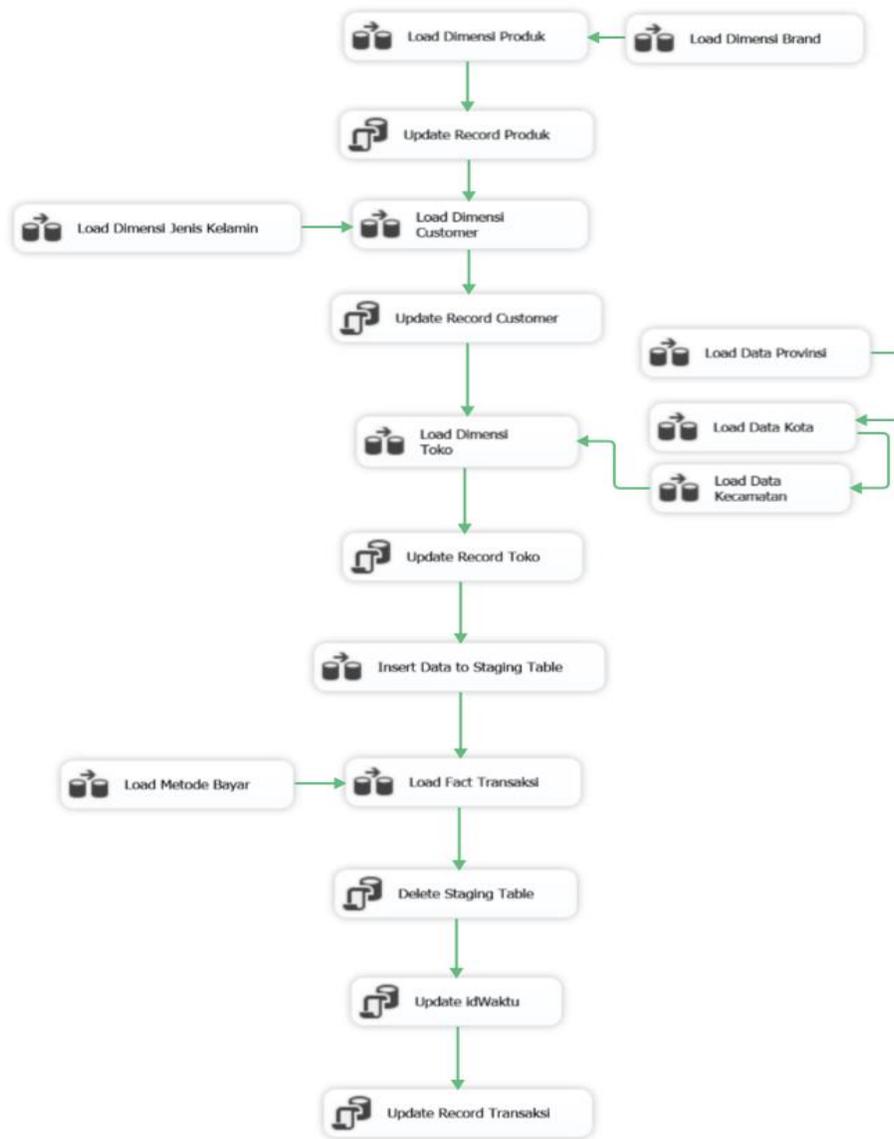
Tahapan dalam implementasi data warehouse dibagi menjadi tiga proses yaitu proses ETL, pembuatan *cube* dan visualisasi.

ETL merupakan proses dimana data dari sumber data (biasanya berupa *database* operasional) diambil untuk kemudian diproses ke dalam data *warehouse*. Setelah menentukan proses bisnis yang akan diambil sebagai acuan data *warehouse* kemudian dilakukan proses *extract, transform & load* data (ETL). Proses ETL menggunakan *Microsoft Server Integration Service*. Proses ini dimulai dengan mempersiapkan sumber data yang diambil dari *database* operasional perusahaan. Proses ETL dilakukan per tabel dengan data dari setiap tabel sumber data akan di ekstrak, disanitasi (dibersihkan dari potensi data tidak lengkap, kesalahan penulisan, data kosong) kemudian di-load ke *database* tujuan untuk disimpan.



Gambar 2. Skema Proses Implementasi Data Warehouse

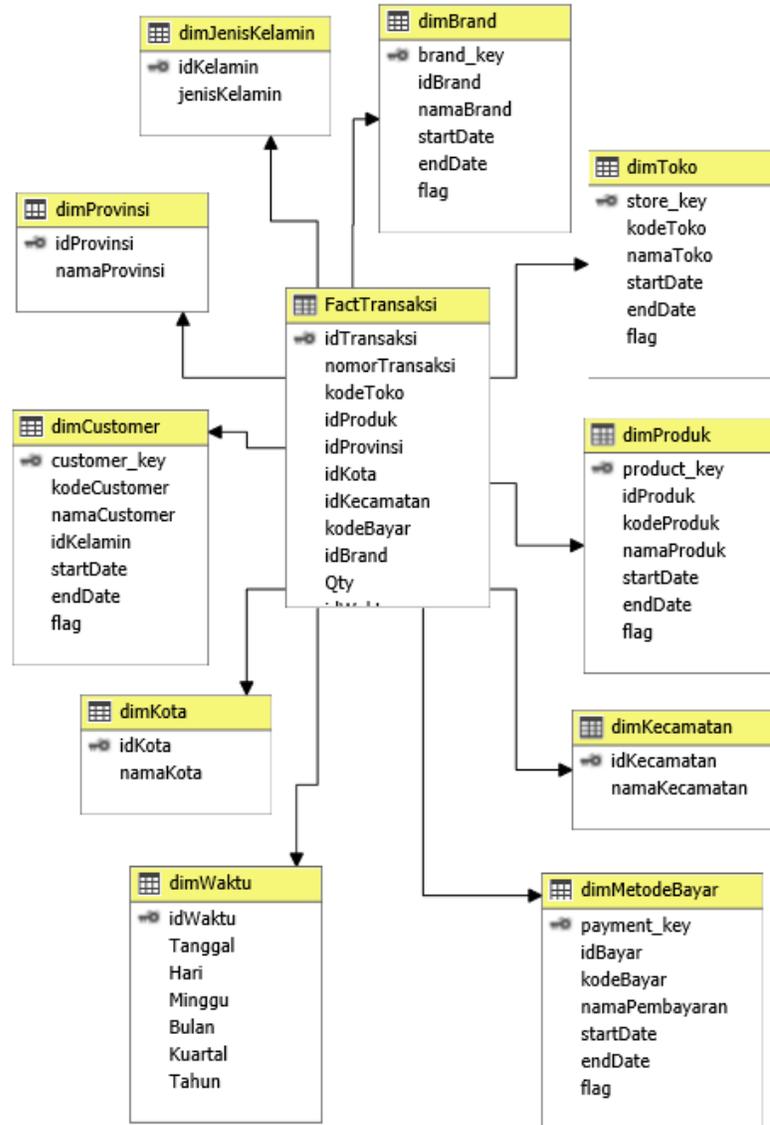
Proses ETL dilakukan pada data terbaru atau data yang belum ada didalam *database* Warehouse. Guna memastikan data yang dimasukkan belum ada pada *database* Warehouse dilakukan perbandingan pada *timestamp* antara data yang ada pada *database* operasional dengan *database* Warehouse. Setelah data terbaru di-load ke dalam *database* Warehouse sebuah *SQL script* akan melakukan *update* pada tabel pada *database* Warehouse yang mengindikasikan bahwa sebuah tabel pada *database* operasional sudah di-copy ke dalam database Warehouse. Proses ETL akan menghasilkan diagram proses ETL.



Gambar 3. Diagram Proses ETL Data Warehouse

Proses *extract* dilakukan terhadap tabel-tabel yang terkait dengan transaksi seperti tabel produk, tabel *brand*, tabel jenis kelamin, tabel *customer*, tabel provinsi, tabel kota, tabel kecamatan, tabel toko dan tabel metode bayar. Setelah proses *extract* disetiap tabel selesai, *timestamp* proses *extract* terakhir akan diupdate. Proses dilanjutkan dengan proses *transform*. Data terkait transaksi di-*transform* di tabel *staging* data. Data dari tabel *staging* kemudian di-*load* ke tabel fakta. Setelah proses *load* selesai data di tabel *staging* akan dihapus dan record proses ETL akan diupdate.

Proses pembuatan *cube* dilakukan menggunakan *Microsoft Server Analysis Service*. Sumber data untuk pembuatan *cube* diambil dari database Warehouse yang telah didapatkan pada proses ETL. Proses ini menghasilkan *OLAP cube* beserta diagram *star flake*.



Gambar 4. Hasil OLAP Cube

Diagram *star flake* dari proses pembuatan *cube* terdiri dari tabel *FactTransaksi* yang memiliki relasi dengan tabel dimensi. Relasi antara tabel *FactTransaksi* dengan tabel dimensi menggunakan *natural key* untuk tabel yang sering mengalami penambahan data secara cepat. Tujuan lain dari penggunaan *natural key* agar setiap update pada tabel dapat ditelusuri *history* tanpa menghapus data di tabel *FactTransaksi* sehingga dapat mendukung *slow changing dimension type 2*. Pada *slow changing dimension type 2* data yang terdapat perubahan akan di-update dengan menyimpan *history update* pada row yang lain. *Slow changing dimension type 2* memungkinkan *history* terkait transaksi dan perubahan pada tabel dapat dilacak sehingga memungkinkan untuk dianalisa dikemudian hari jika diperlukan.

Tabel 1. Performa Query

Nama Query	Query	Proses Query
Transaksi per Bulan	SELECT [Measures].[Fact Transaksi Count] on Columns, NON EMPTY order({[Dim Waktu].[Bulan].[1], [Dim Waktu].[Bulan].[2], [Dim Waktu].[Bulan].[3], [Dim Waktu].[Bulan].[4], [Dim Waktu].[Bulan].[5], [Dim Waktu].[Bulan].[6], [Dim Waktu].[Bulan].[7], [Dim Waktu].[Bulan].[8], [Dim Waktu].[Bulan].[9], [Dim Waktu].[Bulan].[10], [Dim Waktu].[Bulan].[11], [Dim Waktu].[Bulan].[12]} , [Measures].[Bulan], ASC) on Rows FROM [Warehouse];	0.8 detik
Transaksi per Kota	SELECT [Measures].[Fact Transaksi Count] on Columns, NON EMPTY order([Dim Kota].[Nama Kota].members, [Measures].[Dim Kota Count], ASC) on Rows FROM [Warehouse];	1 detik
Transaksi per Produk	SELECT [Measures].[Fact Transaksi Count] on Columns, NON EMPTY order([Dim Produk].[Nama Produk].members, [Measures].[Fact Transaksi Count], DESC) on Rows FROM [Warehouse];	1 detik
Transaksi per Jenis Kelamin	SELECT [Measures].[Fact Transaksi Count] on Columns, NON EMPTY order({[Dim Jenis Kelamin].[Jenis Kelamin].[Pria], [Dim Jenis Kelamin].[Jenis Kelamin].[Wanita]}, [Measures].[Fact Transaksi Count], DESC) on Rows FROM [Warehouse];	0.8 detik

Tabel 1 menunjukkan hasil pengujian performa *query* MDX yang dilakukan pada OLAP *cube* Warehouse. Query dilakukan pada 4 jenis transaksi yang akan divisualisasikan yaitu penjualan per bulan, penjualan berdasarkan *gender*, penjualan per kota dan penjualan per produk. Berdasarkan pengujian rata-rata waktu untuk mengeksekusi *query* paling cepat 0.8 detik dan paling lama 1 detik.

Proses visualisasi menggunakan *Tableau Pro 2020*. Visualisasi dilakukan dengan memasukkan OLAP *cube* sebagai sumber data. Data dari *cube* dikumpulkan ke dalam *measurement & dimension* untuk kemudian divisualisasikan berdasarkan data yang dimasukkan. Data divisualisasikan sesuai dengan kebutuhan bisnis pemangku kepentingan.



Gambar 5. Hasil Visualisasi Sales Dengan Tableau Pro 2020

*Dashboard* visualisasi *sales* merepresentasikan penjualan menjadi 4 kategori yaitu penjualan per bulan, penjualan berdasarkan *gender*, penjualan per kota dan penjualan per produk. Pengguna dapat melakukan filter dan *slicing* terhadap OLAP *cube* tanpa harus mengetahui *query* pada *database*. Setiap data yang difilter oleh pengguna

terhadap satu *worksheet* akan secara otomatis mempengaruhi data yang ada didalam *worksheet* lain sehingga memudahkan pengguna dalam melihat progress penjualan berdasarkan kategori tanpa harus merubah *worksheet* lain secara manual.

## 5. KESIMPULAN

Penelitian ini menghasilkan desain data *warehouse* yang dapat digunakan untuk mengekstrak data terkait penjualan dari database operasional ke dalam data *warehouse* agar dapat dianalisis. Hasil lain penelitian ini berupa *dashboard* visual yang digunakan untuk melihat posisi penjualan. *Dashboard* visual membuat pengguna dapat melakukan filter dan *slicing* terhadap OLAP *cube* tanpa harus melakukan *query* ke *database* atau melakukan penghitungan data penjualan secara manual sehingga mempercepat pengguna dalam memahami kondisi penjualan tanpa harus mengetahui detail teknis pada database. Hasil pengujian performa *query* MDX dalam mengakses data di dalam OLAP *cube* menunjukkan rata-rata waktu yang dibutuhkan untuk eksekusi *query* adalah 1 detik.

## 6. SARAN

Penelitian dapat dikembangkan lebih lanjut agar dapat meng-ekstrak data dari berbagai sumber dan visualisasi dengan aplikasi lain selain *Tableau Pro 2020*. Peneliti juga berharap dapat melakukan implementasi data *warehouse* menggunakan metode lain selain metode Kimball seperti metode Inmon dan metode data *vault*.

## DAFTAR PUSTAKA

- [1] L. McGregor. (2010). Improving the quality and speed of decision making. *Journal of Change Management*, pp 344-356. DOI: 10.1080/714042513.
- [2] Romindo, R., Niar, H., Sipayung, R., Julyanthry, J., Yendrianof, D., Faisal, M., Febrianty, F., Jamaludin, J., Ardiana, D. P. Y., Simarmata, J., Purba, B. (2020). *Sistem Informasi Bisnis*. Medan: Yayasan Kita Menulis.
- [3] Brynjolfsson, Erik & Kristina McElheran. 2019. *Data in Action: Data-Driven Decision Making and Predictive Analytics in U.S. Manufacturing*. Toronto: Rotman School of Management.
- [4] Kubina, Milan & Varmus, Michal & Kubinova, Irena. (2015). Use of Big Data for Competitive Advantage of Company. *Procedia Economics and Finance*. 26. 561-565. 10.1016/S2212-5671(15)00955-7.
- [5] Laudon, K. C. dan Jane P. Laudon. (2010). *Management Information Systems: Managing the Digital Firm, 11<sup>th</sup> Edition*. New Jersey: Prentice Hall.
- [6] Stair, R. M., & G. W. Reynolds. (2010). *Principles of Information Systems, a managerial Approach, 9<sup>th</sup> Edition*. Massachusetts: Course Technology.
- [7] Neha & Garg, K. (2017). Efficient Information Retrieval Using Multidimensional OLAP Cube. *International Research Journal of Engineering and Technology*, vol. 4.
- [8] Yunandar, R., Amir, A., & Rizal, K. (2020). Perancangan Data Warehouse Untuk Informasi Strategi Studi Kasus Penerimaan Siswa Baru STIE Binaniaga Bogor. *Jurnal Teknik Komputer*. 6. 111-120. 10.31294/jtk.v6i1.6861.
- [9] Prasetya P. Hendro. (2016). Visualisasi Informasi Data Perguruan Tinggi Dengan Data Warehouse dan Dashboard System, *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 2, no. 3, doi:10.28932/jutisi.v2i3.504.