**LAPORAN PROYEK MATA KULIAH**

**12S3202 – GUDANG DATA DAN KECERDASAN BISNIS**

**Pembuatan Data Mart Dan Dashboard Pada Spotify**



**Disusun oleh:**

1. 12S18006 – Mei Kristina Panjaitan
2. 12S18023 – Agustina Silaen
3. 12S18055 – Raja Gading Muda Tulen Sihite

**FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO**

**INSTITUT TEKNOLOGI DEL**

**2021/2022**

# DAFTAR ISI

[DAFTAR ISI i](#_heading=h.gjdgxs)

[1.](#_heading=h.30j0zll) Pendahuluan 1

[2.](#_heading=h.3znysh7) Analisis 2

[3.](#_heading=h.2et92p0) Desain 3

[4.](#_heading=h.1t3h5sf) Implementasi 5

[5.](#_heading=h.17dp8vu) Penutup 7

[LAMPIRAN 8](#_heading=h.3rdcrjn)

# Pendahuluan

Pada Bab ini bertujuan untuk mengidentifikasi masalah dalam penelitian dalam Latar belakang dan ruang lingkup, tujuan tugas besar dari mata kuliah Data Warehouse and Business Intelligence.

* 1. **Latar Belakang**

Dalam Dokumen ini akan berisikan dengan Data Mart dan Dashboard pada spotify, dengan itu untuk melakukan proses dalam pembuatan data mart dan dashboard dilakukan penyajian data dengan menganalisis pembuatan data mart. Proses yang dilakukan dengan mengumpulkan sumber data staging dengan tujuan untuk mengelola, menggabung data agar tidak terjadi kesalahan pada data.

Setelah melakukan analisis pada data maka dilakukan pembuatan data mart untuk memenuhi tujuan dan manfaat lalu kemudian akan dilakukan pembuatan dashboard, untuk mengambil keputusan.

* 1. **Ruang Lingkup**

Data warehouse merupakan bagian dari Decision Support System. Data warehouse adalah subject oriented, integrated, time variant, dan nonvolatile dari koleksi data untuk membantu proses pembuatan keputusan [1]

1. Pembangunan Data warehouse

Pembangunan data warehouse ada beberapa pendekatan, salah satunya yaitu pendekatan Down approach yang didukung oleh inmon dan bottom Up approach yang didukung oleh kimball & Margy Ross

1. Top- Down Approach

Data warehouse menyimpan data transaksi yang atomik yang diekstrak dari satu atau beberapa sumber dan diintegrasikan dalam enterprise data model yang dinormalisasikan. Dari enterprise data model inilah data disimpulkan, dipecahkan berdasarkan dimensinya dan didistribusikan ke satu atau beberapa dependen data mart

1. Bottom-Up approach

pendekatan bottom-up dimulai dengan eksperimen dan prototipe. ini berguna dalam tahap awal pemodelan bisnis dan pengembangan teknologi. pembuatan enterprise data warehouse membutuhkan waktu dan biaya yang terlalu lama, sedangkan membuat data mart yang saling terisolasi satu sama lain membuat sistem yang dihasilkan menjadi tidak terintegrasi.

2. Komponen Data warehouse

Secara umum sebuah sistem data warehouse mempunyai 4 komponen terpisah yang mempunyai peran dan fungsi yang spesifik, yaitu operational source system, data staging area, data presentation area dan data access tool

1. Operational source System

Merupakan sistem operasional yang menyimpan transaksi dari proses bisnis. Prioritas utama dari sistem ini adalah kinerja proses data dan ketersediaannya.

1. Data Staging Area

Merupakan area penyimpanan dan serangkaian proses yang biasa disebut sebagai ExtractTransformation-load (ETL) terhadap data sebelum digunakan dalam data warehouse

1. Data Presentation Area

Data preparation area dapat berupa data mart data mart yang terintegrasi. Dimana data mart adalah suatu bagian logik dari keseluruhan suatu data warehouse yang spesifik pada satu proses bisnis dan diorganisasikan serta ditargetkan untuk satu bagian dalam perusahaan

1. Data access Tool

Data access tool dapat berupa aplikasi ad hoc query yang sederhana atau aplikasi data mining yang kompleks.

3. Dashboard

Dashboard merupakan tampilan visual mengenai informasi paling penting yang diperlukan untuk mencapai satu tujuan atau lebih dan dapat diatur di satu layar sehingga lebih mudah dipantau oleh user. Tujuan dari pembuatan dashboard ini yaitu untuk memantau informasi, beberapa pekerjaan memerlukan dashboard untuk dipantau secara real time karena aktivitas yang dipantau saat ini sedang terjadi.

4. Pentaho data integration (PDI)

Pentaho data Integration (PDI) adalah software dari Pentaho yang dapat digunakan untuk proses ETL data. PDI dapat digunakan untuk migrasi data, membersihkan data, loading dari file ke database atau sebaliknya dalam volume besar. PDI menyediakan graphical user interface dan drag-drop komponen yang memudahkan user.

5. Power BI

Power BI merupakan aplikasi Microsoft yang dapat mengunggah data dan membangun dashboard. Aplikasi ini mampu untuk menampilkan data secara realtime dalam bentuk dashboard yang mampu melihat detail data secara lebih lengkap. Power BI dapat memvisualisasikan data dari Excel, SQL server, web, CSV.

* 1. **Tim Pengembang**

Pada Tabel 1 disajikan susunan tim pengembang pada proyek ini.

**Tabel 1. Susunan Tim Pengembang**

| No. | Nama | Peran | Tanggung Jawab |
| --- | --- | --- | --- |
|  | Mei Kristina Panjaitan | Analyst | Ketua |
|  | Agustina Silaen | Analyst | Anggota |
|  | Raja Gading Muda Tulen Sihite | Analyst | Anggota |

# Analisis

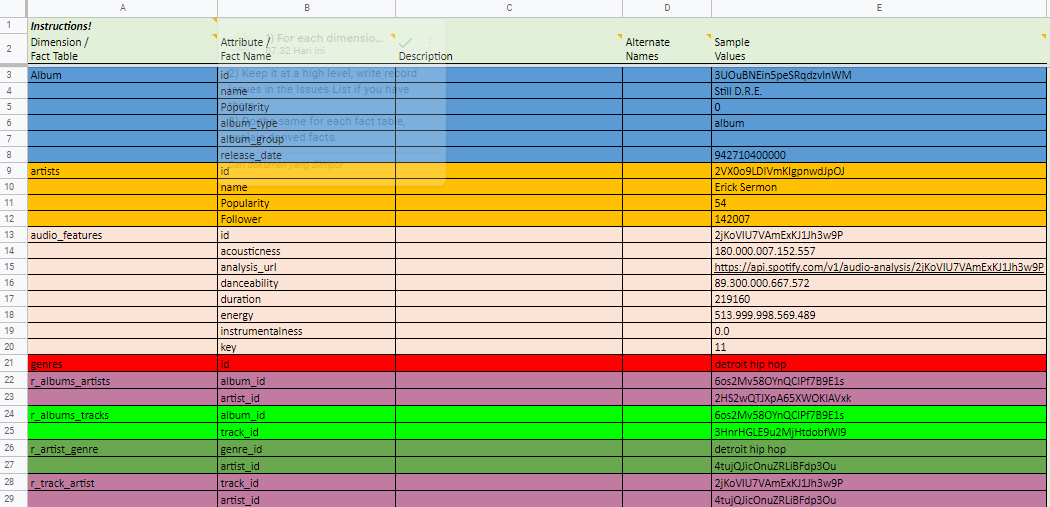
Dalam membangun data warehouse, diperlukan perancangan sistem terintegrasi yang akan menghubungkan data dari sumber hingga dapat disajikan dalam bentuk visualisasi data sehingga akan dihasilkan informasi yang dibutuhkan. Dengan adanya data warehouse pada data spotify yang sudah disediakan.

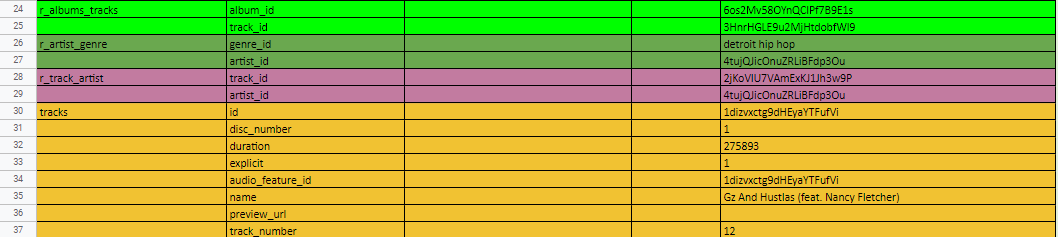
* 1. **Bisnis Proses**

Dalam sub bab ini berisikan bisnis proses pada proyek yang akan kami kembangkan untuk melakukan desain data mart dan melakukan integrasi pada proses ETL,bisnis proses yang ingin dikembangkan adalah rating Movie

* 1. **Atribut dan Metrik**

Dalam sub bab ini akan berisikan attribut dan metrik yang berhubungan dengan proyek





* 1. **Isu**

Pada bagian ini berisikan kebutuhan informasi yaitu melakukan analisa data, dan penentuan data, dan informasi. Tahap ini sangat penting karena akan mempertimbangkan tahapan desain pada dataset, dan

# Desain

Bagian ini berisi *detailed bus matrix*, *detailed dimensional modeling*, *dimensional hierarchies*, *dimensional data model*, spesifikasi *extract, transform, load* (ETL), *detailed ETL flow for each source to target*, dan *user interface dashboard mockup*.

* 1. ***Detailed Bus Matrix***

<Bagian ini mengacu pada [High-Level-Dimensional-Modeling-Workbook.xlsx](https://apipuro.del.ac.id/v1/file/7a0717c21c07474d85483228d3b15a20).>

* 1. ***Detailed Dimensional Modeling***

<Bagian ini mengacu pada [Detailed-Dimensional-Modeling-Workbook-KimballU.xlsm](https://apipuro.del.ac.id/v1/file/e11033a04e21c267f7babf6be10e4504).>

* 1. ***Dimensional Hierarchies***

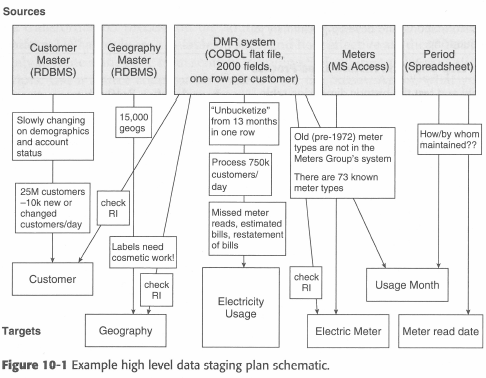
<Gambarkan hirarki pada setiap tabel dimensi.>

* 1. ***Dimensional Data Model***

<Gambarkan *schema* model dimensional.>

* 1. ***High-Level Source to Target Map***

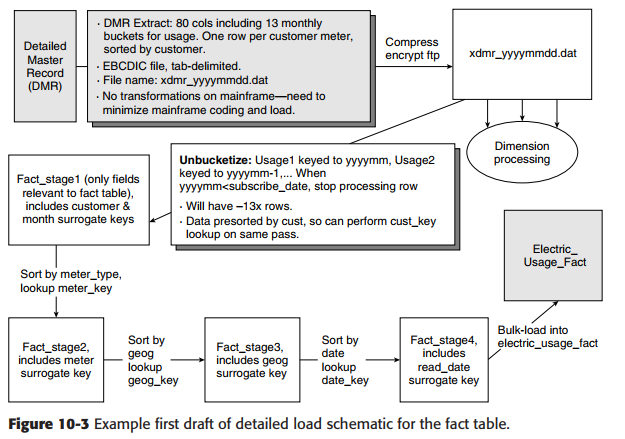
<Gambarkan *high-level source to target map*. Contoh dapat dilihat pada Gambar 1.>



**Gambar 1. Diagram Tingkat Tinggi Sumber ke Target**

* 1. ***Detailed ETL Flow for Each Source to Target***

<Tuliskan aliran ETL secara rinci untuk setiap aliran sumber ke target. Contoh dapat dilihat pada Gambar 2.>



**Gambar 2. Diagram Rinci ETL DMR ke Tabel Electric\_Usage\_Fact**

* 1. ***Business Intelligence Front End Mockup***

<Gambarkan *mockup BI front end*, misalnya *dashboard* berserta panel-panel yang sesuai dengan bisnis proses yang telah didefinisikan.>

# Implementasi

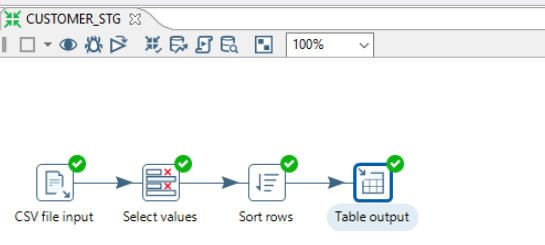
Bagian ini berisi ROLAP *schema*, kode program pada ETL, basis data MOLAP, dan *Business Intelligence Front End*.

* 1. ***ROLAP Schema***

<Tampilkan cuplikan ROLAP *schema* yang menunjukkan *fact tables*, *dimension tables*, *star/snowflake/galaxy schema diagram*, dan PK/FK.>

* 1. ***ETL***

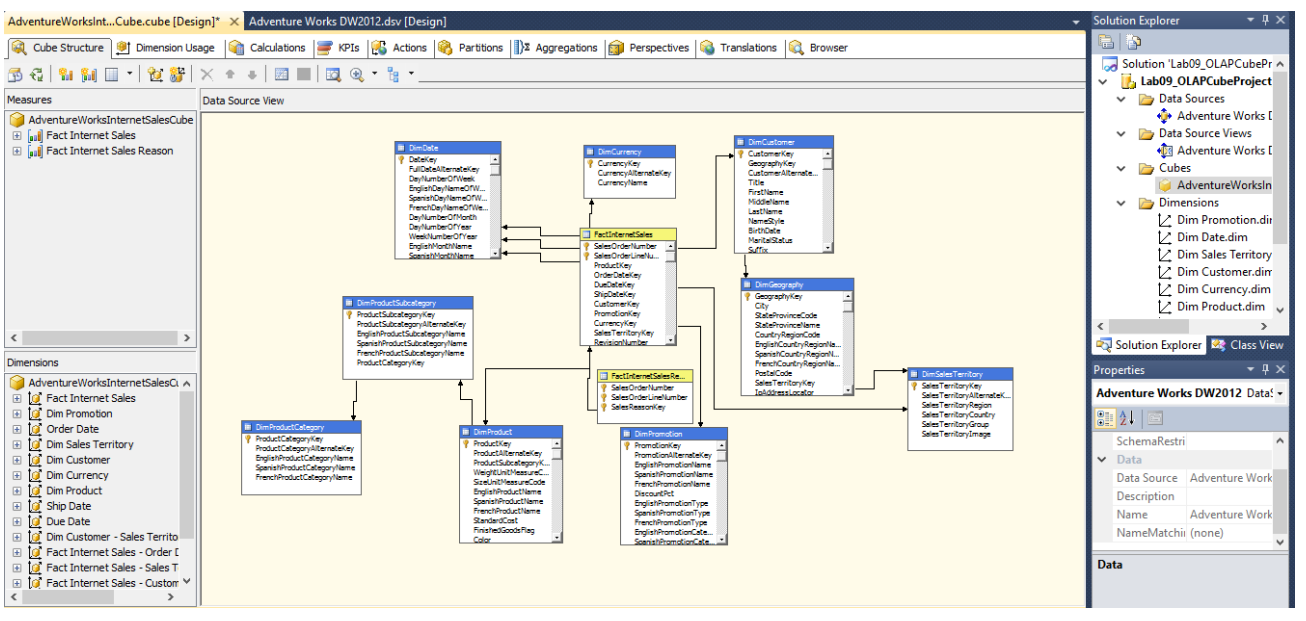
<Tuliskan kode program (atau diagram proses apabila menggunakan **SQL Server Integration Services** ataupun **Spoon**) yang terlibat dalam proses ETL. Misalnya, kode program untuk *populates ROLAP schema from sources*, kode program untuk *initial load*, dan kode program untuk *subsequent loads*. Contoh dapat dilihat pada Gambar 3.>



**Gambar 3. Cuplikan ETL dari Sumber ke Tabel Staging Customer**

* 1. ***Analysis Services MOLAP Database***

<*Cubes to match star schemas / business processes, dimensional hierarchies configured, facts / KPI’s as appropriate.* Contoh dapat dilihat pada Gambar 4.>



**Gambar 4. Cuplikan Internet Sales Cube**

* 1. ***Dashboard***

<Tampilkan cuplikan *dashboard* serta beri penjelasan penggunaannya seperti apa.>

# Penutup

* 1. **Kesimpulan**

<Tuliskan apakah sistem selesai dibangun dan apakah dapat menyelesaikan masalah yang ditulis pada latar belakang.>

* 1. **Saran**

Tuliskan saran pengembangan.>

# LAMPIRAN

Referensi

[1] W. H. Inmon, Building the Data Warehouse, Third. Canada: Robert Elliott, 2002.