**LAPORAN PROYEK MATA KULIAH**

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

**SPOTIFY**



**Disusun oleh:**

12S19001 – Raynaldo Silalahi

**FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO**

**INSTITUT TEKNOLOGI DEL**

**April 2022**

# DAFTAR ISI

[DAFTAR ISI i](#_Toc100835754)

[1. Pendahuluan 1](#_Toc100835755)

[2. Analisis 2](#_Toc100835756)

[3. Desain 3](#_Toc100835757)

[4. Implementasi 5](#_Toc100835758)

[5. Penutup 7](#_Toc100835759)

[LAMPIRAN 8](#_Toc100835760)

# Pendahuluan

Bagian ini berisi latar belakang dan tujuan pengerjaan proyek.

* 1. Latar Belakang

Perkembangan dan kemajuan teknologi informasi selalu dituntut untuk dapat dimanfaatkan sebagai penunjang proses bisnis dalam sebagian besar instansi. Teknologi informasi dituntut untuk mendapat memberikan kemudahan bagi manusia dalam berkativitas seperti dalam pengambilan keputusan, akses informasi, memecahkan masalah komputasi, pengiriman data, dan lain sebagainya.

Data staging merupakan salah satu proses atau lokasi dimana data dari system sumber disalin, yang terdapat pada data warehouse. Kemudian data staging dibutuhkan sebagai tempat dalam melakukan proses menyeleksi, mengumpulkan, dan menyiapkan data yang diberikan untuk dimuat ke dalam data warehouse sehingga dapat dipresentasikan dalam bentuk laporan reporting. Jadi data staging bertujuan untuk mengumpulkan data yang terpisah-pisah menjadi satu tempat/ lokasi.

Data mart merupakan bagian dari data warehouse yang berada pada tingkatan yang lebih kecil seperti level departemen pada suatu organisasi atau perusahaan. Data mart tersebut hanya berfokus menangani sebuah proses bisnis, misalkan pada proses bisnis Top Genre dalam Spotify.

Permasalahan dalam Top Genre pada Spotify yaitu belum adanya data mart yang dibutuhkan untuk penyimpanan data yang nantinya akan dibutuhkan oleh perusahaan Spotify.

* 1. Ruang Lingkup

Ruang lingkup adalah suatu hal yang dibutuhkan dalam pengembangan system data mart pada Spotify. Dalam pengembangan ini telah disediakan sumber data dari pihak ketiga yaitu berupa SQLite. Kemudian pendekatan pengembangan system menggunakan konsep Kimball Approach untuk data warehouse dan konsep Inmon Approach untuk proses bisnis intelligent.

Perangkat pendukung

|  |  |
| --- | --- |
| Nama | Keterangan |
| Perangkat Lunak Basis Data | Microsoft SQL Server, PostgreSQL |
| Perangkat Lunak ETL | Excel |
| Perangkat Lunak Dashboard | Tableau |

* 1. Tim Pengembang

# Analisis

Bagian ini berisi bisnis proses, *bus matrix*, *attributes and metric list*, *issues list*.

* 1. Bisnis Proses

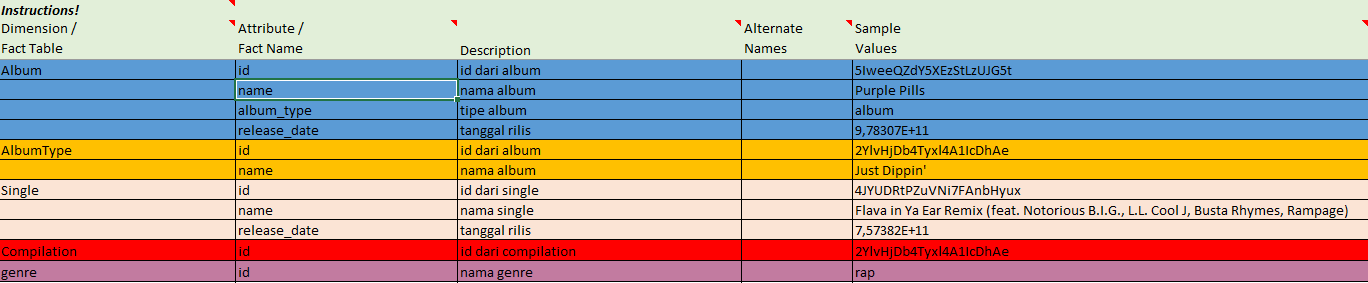
Dalam pembangunan system data mart dan dashboard pada Spotify, proses bisnis yang nantinya akan dikembangkan yaitu :

* Top genre

Team ingin memeriksa jumlah genre yang paling sering digunakan dalam sebuah lagu berdasarkan album, single, genre, compilation pada dataset.

* 1. Atribut dan Metrik

Daftar atribut dan metrik (*fact*) yang Anda gunakan dalam sistem yang akan dibangun dapat dilihat pada gambar berikut ini:



* Dataset

Proyek ini menggunakan Dataset Playslist yang dirilis Malte Grosse dan dapat diunduh melalui situs Kaggle. Dataset ini terdiri dari jutaan list Lagu dan terdapat metadata berupa album, jumlah genre, dan lain lain.

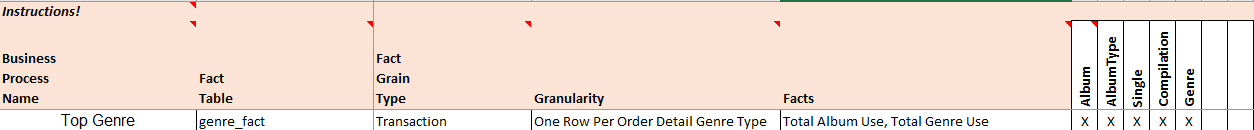
* 1. Isu

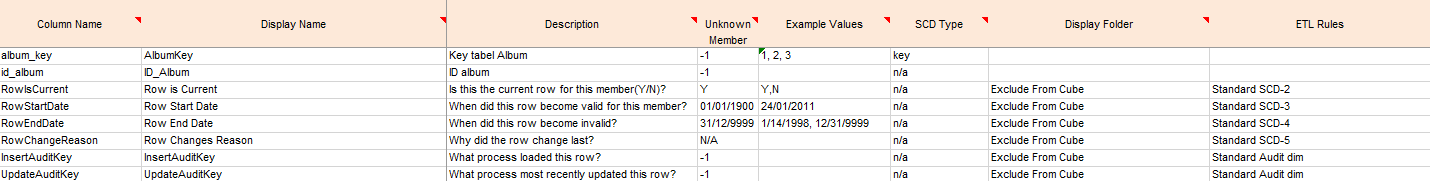
Masalah yang ditemui pada sistem data yang perlu di pertimbangkan dalam tahap desain data mart secara detail dan tahap desain sistem extract, transform, load :

# Desain

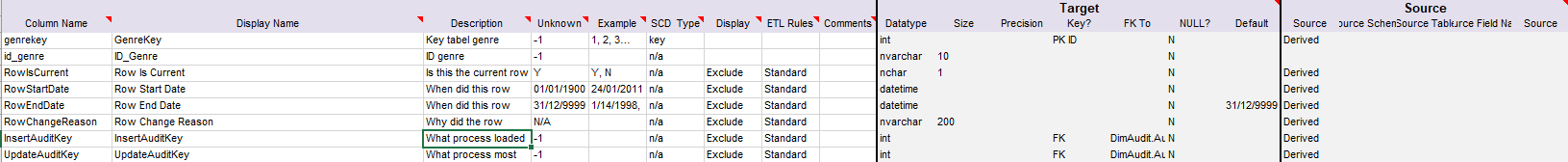
Bagian ini berisi *detailed bus matrix*, *detailed dimensional modeling*, *dimensional hierarchies*, *dimensional data model*, spesifikasi *extract, transform, load* (ETL), *detailed ETL flow for each source to target*, dan *user interface dashboard mockup*.

* 1. Detailed Bus Matrix





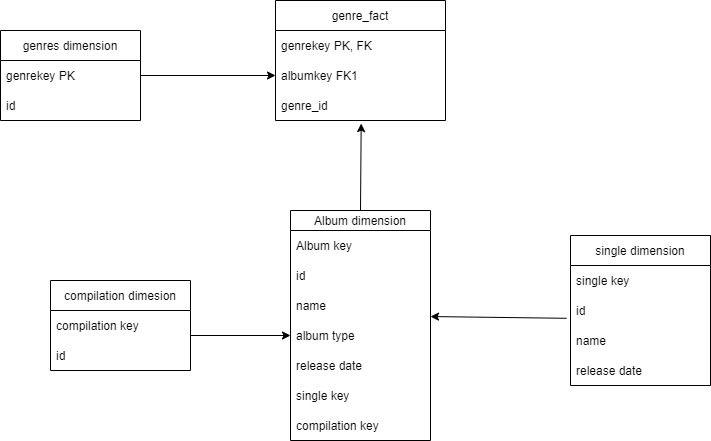




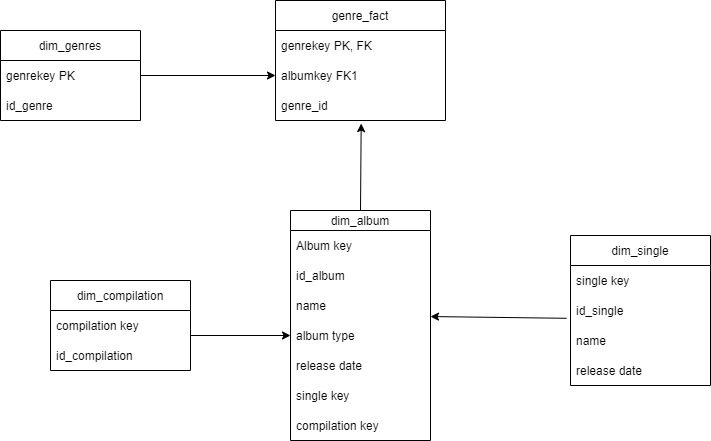
* 1. Detailed Dimensional Modeling



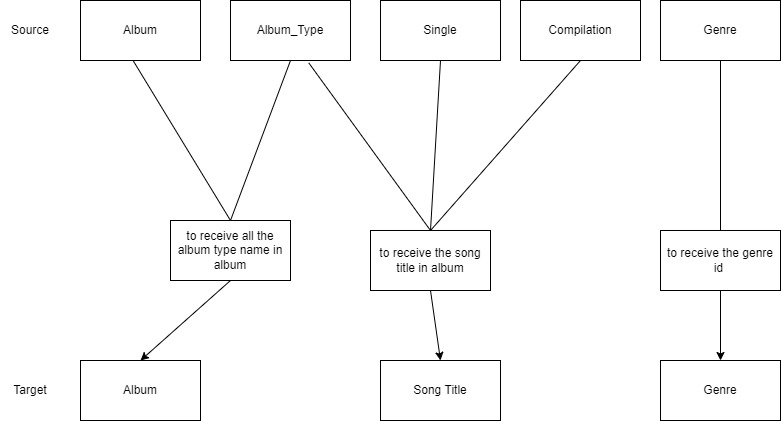
* 1. Dimensional Hierarchies



* 1. Dimensional Data Model



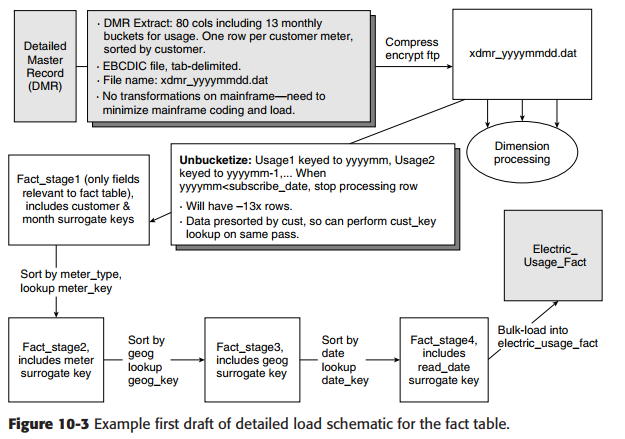
* 1. High-Level Source to Target Map



Gambar . 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 . 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

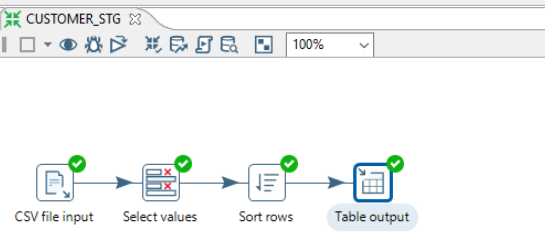
Bagian ini berisi ROLAP *schema*, kode program pada ETL, basis data MOLAP, dan *Business Intelligence Front End*.

* 1. ROLAP Schema

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

* 1. ETL

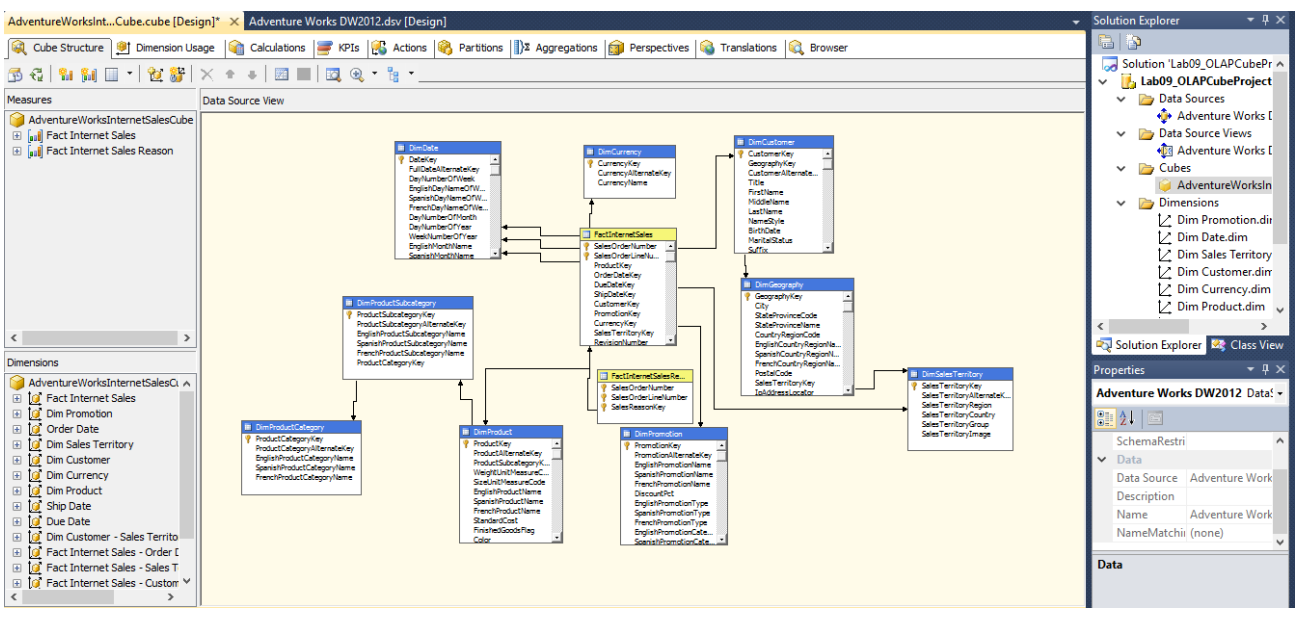
<Tuliskan kode program (atau diagram proses apabila menggunakan **SQL Server Integration Services** ataupun **Spoon**) yang terlibat dalam proses ETL. Misalnya, kode program untuk *populates ROLAP schema from sources*, kode program untuk *initial load*, dan kode program untuk *subsequent loads*. Contoh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



Gambar 3. Cuplikan ETL dari Sumber ke Tabel Staging Customer

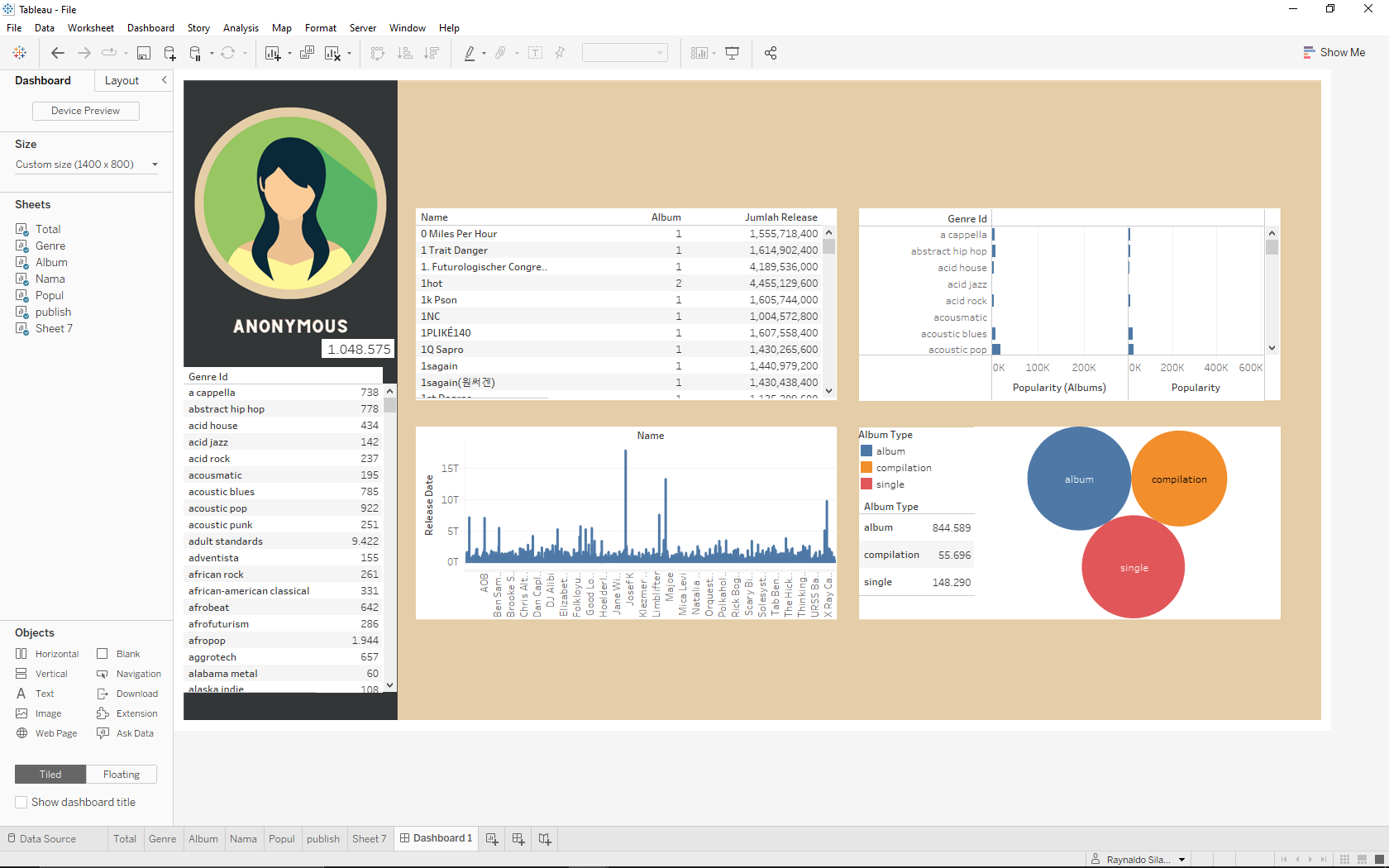
* 1. Analysis Services MOLAP Database

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



Gambar . Cuplikan Internet Sales Cube

* 1. Dashboard



<https://prod-apnortheast-a.online.tableau.com/#/site/hardwarecom/views/Spotify/Total?:origin=card_share_link&:embed=n>

# Penutup

* 1. Kesimpulan

Belum selesai

* 1. Saran

Tuliskan saran pengembangan.>

# LAMPIRAN

<Opsional.>