

LAPORAN PROYEK MATA KULIAH
12S3202 – GUDANG DATA DAN KECERDASAN BISNIS

Pembuatan Sistem Data Mart dan Dashboard
pada Spotify



Disusun oleh:

12S18001	Cindy Angelia Siregar
12S18014	Giovanni Situmorang
12S18017	Putri Yohana Panjaitan
12S18049	Natasya Sitorus
12S18061	Angela Friscilia Simamora

FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO
INSTITUT TEKNOLOGI DEL

Mei 2022

DAFTAR ISI

DAFTAR ISI.....	i
DAFTAR TABEL.....	iii
DAFTAR GAMBAR	iv
1. Pendahuluan.....	1
1.1. Latar Belakang	1
1.2. Ruang Lingkup.....	2
1.3. Tim Pengembang.....	2
2. Analisis	3
2.1. Bisnis Proses	3
2.2. Atribut dan Metrik.....	3
2.3. Isu	4
3. Desain	5
3.1. Detailed Bus Matrix	5
3.2. Detailed Dimensional Modeling	5
3.2.1 Dimensi Album.....	5
3.2.2 Dimensi Artis	6
3.2.3 Dimensi Fitur Audio	6
3.2.4 Dimensi Genre	6
3.2.5 Dimensi Track.....	7
3.3. Dimensional Hierarchies	7
3.3.1 Dimensi Album.....	7
3.3.2 Dimensi Artis	8
3.3.3 Dimensi Fitur Audio	8
3.3.4 Dimensi Genre	8
3.3.5 Dimensi Track.....	8
3.4. Dimensional Data Model	8
3.5. High-Level Source to Target Map	9
3.6. Detailed ETL Flow for Each Source to Target	9
3.6.1 Dimensi Artis	9
3.6.2 Dimensi Album.....	10
3.6.3 Dimensi Fitur Audio	10

3.6.4	Dimensi Genre	11
3.6.5	Dimensi Track.....	11
3.7.	Business Intelligence Front End Mockup	12
4.	Implementasi.....	13
4.1.	ETL.....	13
4.1.1	Staging	13
4.1.2	Dimension	15
4.2.	Dashboard.....	19
5.	Penutup	20
5.1.	Kesimpulan.....	20
5.2.	Saran.....	20
LAMPIRAN.....		21

DAFTAR TABEL

Tabel 1. Susunan Tim Pengembang.....	2
Tabel 2 Atribut dan Metrik	3
Tabel 3 Isu.....	4

DAFTAR GAMBAR

Gambar 1 Detailed Bus Matrix	5
Gambar 2 Detailed Dimensional Modeling Dimensi Album.....	5
Gambar 3 Detailed Dimensional Modeling Dimensi Artis.....	6
Gambar 4 Detailed Dimensional Modeling Dimensi Fitur Audio	6
Gambar 5 Detailed Dimensional Modeling Dimensi Genre	6
Gambar 6 Detailed Dimensional Modeling Dimensi Track	7
Gambar 7 Dimensional Hierarchi Album	7
Gambar 8 Dimensional Hierarchi Artis	8
Gambar 9 High Level Source to Target Map.....	9
Gambar 10 Detailed ETL Flow Dimensi Artis	9
Gambar 11 Detailed ETL Flow Dimensi Album	10
Gambar 12 Detailed ETL Flow Dimensi Fitur Audio	10
Gambar 13 Detailed ETL Flow Dimensi Genre	11
Gambar 14 Detailed ETL Flow Dimensi Track.....	11
Gambar 15 Business Intelligence Front End Mockup	12
Gambar 16 Staging Album	13
Gambar 17 Staging Artis	13
Gambar 18 Staging Fitur Audio.....	14
Gambar 19 Staging Track	14
Gambar 20 Staging Genre.....	15
Gambar 21 Dimension Album.....	15
Gambar 22 Dimension Artis	16
Gambar 23 Dimension Audio	16
Gambar 24 Dimension Genre	17
Gambar 25 Dimension Track.....	17
Gambar 26 Staging Value.....	18
Gambar 27 Pre Fact	18
Gambar 28 Proses ETL.....	19

1. Pendahuluan

Bab pendahuluan berisi penjelasan terkait latar belakang pemilihan topik proyek, ruang lingkup proyek dan tim pengembang dalam proyek.

1.1. Latar Belakang

Meningkatnya jumlah interaksi sosial publik terjadi melalui komputer, dimana semuanya dibagikan mulai dari lelucon, cerita, berita dan terutama dunia industri musik. Salah satu layanan aplikasi media sosial yang menawarkan musik kepada pengguna adalah *spotify*. *Spotify* adalah salah satu layanan musik digital, podcast, dan video yang memberikan akses jutaan lagu dari kreator di seluruh dunia. *Spotify* menawarkan musik kepada pengguna, dimana pengguna dapat mencari musik berdasarkan artis, album, atau genre serta pengguna dapat mengedit dan berbagi daftar putar.

Spotify juga memiliki beberapa fitur salah satunya adalah dapat menambah lagu sebanyak mungkin dan juga dari lagu-lagu yang bertambah, maka artis atau penyanyi juga akan bertambah. Akan tetapi seiring perkembangan lagu, maka data mengenai lagu juga semakin banyak dan biasanya *spotify* akan menyimpan lagu dan artis tersebut dalam waktu yang lama. Kumpulan data yang begitu banyak mengandung informasi mengenai lagu dan juga artis atau penyanyi yang dapat dianalisis. Saat ini data yang digunakan untuk merekapitulasi data *spotify* sudah ada, akan tetapi karena banyaknya fitur yang terdapat pada *spotify*, maka diperlukan sistem yang akan mengintegrasikan dan dapat merangkum data-data tersebut.

Salah satu pendekatan yang dapat digunakan dalam proses mengintegrasikan data adalah dengan membuat *data mart*. Dalam mengintegrasikan data, proses yang akan dilakukan adalah mengumpulkan data dari sumber dan diolah melalui proses ETL, dan setelah itu akan data akan digabungkan dalam satu penyimpanan. Data yang sudah diintegrasikan ke dalam satu penyimpanan akan ditampilkan dalam bentuk *dashboard* yang bertujuan untuk mempermudah proses pemantauan. *Dashboard* merupakan salah satu penerapan kecerdasan bisnis, yang digunakan sebagai sistem yang menampilkan informasi visual yang mudah dipahami oleh pengguna berdasarkan data-data yang sudah dikumpulkan dan diolah.

Pada pengerjaan proyek ini berfokus dalam pengerjaan menyaring informasi popularitas artis atau penyanyi pada setiap lagu di *Spotify* yang dapat difilter setiap minggu ataupun setiap bulan. Banyaknya jumlah musik yang dipublikasikan oleh para musisi dunia juga membuat pengelolaan terhadap basis data sangat diperlukan dalam menentukan kepopuleran artisnya. Oleh karena itu, pada pengerjaan proyek ini dilakukan analisis dengan membuat sistem data mart dan dashboard untuk *Spotify* supaya setiap informasi dapat diintegrasikan dengan baik.

1.2. Ruang Lingkup

Adapun ruang lingkup pada proyek ini diantaranya adalah.

1. Dataset yang digunakan adalah dataset *spotify.sqlite*.
2. Alat ETL yang digunakan pada proyek ini adalah *pentaho (spoon)* untuk melakukan proses ETL dan *PostgreSQL* untuk DBMS.

1.3. Tim Pengembang

Pada Tabel 1 disajikan susunan tim pengembang pada proyek ini.

Tabel 1. Susunan Tim Pengembang

No.	Nama	Peran	Tanggung Jawab
1	Cindy Angelia Siregar	Ketua	Desain
2	Giovanni Situmorang	Anggota	Implementasi
3	Putri Yohana Panjaitan	Anggota	Implementasi
4	Natasya Sitorus	Anggota	Desain
5	Angela Friscilia Simamora	Anggota	Desain

2. Analisis

Bab analisis berisi penjelasan terkait bisnis proses, *bus matrix*, daftar atribut dan metrik, dan isu pada proyek.

2.1. Bisnis Proses

Pada proyek ini proses bisnis yang akan dibangun adalah pelaporan artis terpopuler. Dimana proses bisnis ini bertujuan untuk menyaring informasi popularitas artis atau penyanyi pada setiap lagu di *Spotify* yang dapat difilter setiap minggu ataupun setiap bulan

2.2. Atribut dan Metrik

Pada bagian ini akan ditampilkan daftar dimensi dan atribut yang dipakai pada proyek ini yaitu yang berhubungan dengan pelaporan artis terpopuler. Dimensi yang digunakan adalah *albums*, *artist*, *audio_features*, *genres*, *tracks*.

Tabel 2 Atribut dan Metrik

Dimension Table	Attribute/Fact Name	Description
album	id	id dari album
	name	nama album yang digunakan
	album_group	group dari album
	album_type	tipe album yang digunakan
	release_date	tanggal rilis album
	popularity	popularitas dari album
artis	id	id artis
	name	nama dari artis
	popularity	popularitas artis
	followers	pengikut artis
fitur audio	id	id dari fitur audio
	acousticness	fitur akustik dalam audio
	analysis_url	analisis link url untuk fitur audio
	danceability	genre musik dance untuk fitur audio
	duration	genre musik dance untuk fitur audio
	energy	kekuatan audio fitur
	instrumentalness	instrumental yang ada pada audio
	key	kunci fitur audio
	liveness	pendeteksi hidup atau mati dari fitur audio
	loudness	kuat dari audio
	mode	mode yang digunakan dalam audio
	speechiness	speechiness dari fitur audio
	tempo	tempo dari audio

Dimension Table	Attribute/Fact Name	Description
	time_signature	waktu untuk menandakan beat audio
	valence	valensi dalam fitur audio
genre	id	Id dari genre
track	id	id dari track
	disc_number	nomor dari disc yang diputar
	duration	durasi dari track
	explicit	kejelasan track yang berjalan
	name	nama track
	preview_url	link url yang sebelumnya diputar
	track_number	nomor dari track
	popularity	popularitas track
	is_playable	keterangan bahwa track dapat dimainkan

2.3. Isu

Tabel 3 Isu

No	Task/Topic	Issue	Reported By	Responsible	Status
1.	Menentukan Topik Proyek	Menentukan bisnis proses apa saja yang dibuat dalam proyek.	Semua anggota	Semua anggota	Solved
2.	<i>Detailed Dimensional Modeling</i>	Menentukan <i>Detailed Dimensional Modeling</i>	Semua anggota	Semua anggota	Solved
3.	ETL process	Menentukan proses ETL setiap dimensional yang ada	Semua anggota	Semua anggota	Solved

3. Desain

Bab desain menjelaskan tentang *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*.

3.1. Detailed Bus Matrix

Pada bagian ini menampilkan *detail bus matrix* yang ada pada proyek, dimana bisnis proses yang digunakan adalah pelaporan artis terpopuler, tabel faktanya adalah artis_terpopuler, *fact grain type* adalah *accumulating snapshot*, kemudian *granularity* pada proses bisnis ini adalah satu baris untuk setiap artis per lagu, per album, per track, per genre, per fitur audio dan faktanya adalah total popularitas, jumlah lagu, jumlah artist dan jumlah album.

Business Process Name	Fact Table	Fact Grain Type	Granularity	Facts	albums	artists	audio_features	genres	tracks
pelaporan artis terpopuler	artis_terpopuler_fact	accumulating snapshot	satu baris untuk setiap artis per lagu, per album, per track, per genre, per fitur audio	Total popularitas, jumlah lagu, jumlah album	x	x	x	x	x

Gambar 1 Detailed Bus Matrix

3.2. Detailed Dimensional Modeling

Pada bagian ini adalah menampilkan *dimensional modeling* dari proses bisnis pelaporan artis terpopuler, dimana terdapat lima dimensi diantaranya adalah *albums*, *artist*, *audio_features*, *genres*, *tracks*.

3.2.1 Dimensi Album

Berikut merupakan detail dari dimensi album yang menjadi sumber data album spotify. Dimensi ini dapat ditemukan pada source db spotify.sqlite dengan keterangan atribut sebagai berikut.

Table Name	DimAlbum	Home Page									
Table Type	Dimension										
Display Name	Album										
Database Schema	spotify.sqlite										
Table Description	tabel album merupakan <i>slowly changing dimension</i> untuk berbagai informasi tentang album berasal dari tabel albums di spotify.sqlite										
Comment											
Biz Filter Logic											
Size	satu untuk setiap album										
Generate Script?	Y										

Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
pk_albums	pk_albums	Kunci utama pengganti	-1	1, 2, 3, ...	key	int			PK ID	N			Derived				
id	id	Kunci bisnis dari sumber data (alias kunci alam)	-1		key	text			N				spotify	dbo	albums	id	text
name	name	nama album		2/koVU7vaMeX9s1 Alkanoliki feat Erik Serannon, J Ro & Tash	2	text			N				spotify	dbo	albums	name	text
album_group	album_group	nama grup album		album	2	text							spotify	dbo	albums	album_group	text
album_type	album_type	nama tipe album		0548336000000	2	text			N				spotify	dbo	albums	album_type	text
release_date	release_date	tanggal rilis album		0	2	bigint			N				spotify	dbo	albums	release_date	bigint
popularity	popularity	popularity album			2	bigint			N	0			spotify	dbo	albums	popularity	bigint

Gambar 2 Detailed Dimensional Modeling Dimensi Album

3.2.2 Dimensi Artis

Berikut merupakan detail dari dimensi artis yang menjadi sumber data album spotify. Dimensi ini dapat ditemukan pada source db spotify.sqlite dengan keterangan atribut sebagai berikut.

Table Name	DimArtis																
Table Type	Dimension	<div>Home Page</div>															
Display Name	Artis																
Database Schema	spotify.sqlite																
Table Description	tabel artis merupakan slowly changing dimension untuk berbagai informasi tentang artis																
Comment	berasal dari tabel artists di spotify.sqlite																
Biz Filter Logic																	
Size	satu untuk setiap data artis																
Generate Script?	Y																
						Target							Source				
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
pk_artis	artis_key	Kunci utama pengganti	-1	1, 2, 3 ...	key	int				PK ID	N		Derived	spotify			
id	id	Kunci bisnis dari sumber data untuk tabel artis (alias kunci alam)	-1	1, 2, 3 ...	key	text					N		spotify	dbo	artists	id	text
name	name	nama dari artis			1	text					N		spotify	dbo	artists	name	text
popularity	popularity	populartias artis			2	bigint					N	0	spotify	dbo	artists	popularity	bigint
followers	followers	pengikut artis			2	bigint					N	0	spotify	dbo	artists	followers	bigint

Gambar 3 Detailed Dimensional Modeling Dimensi Artis

3.2.3 Dimensi Fitur Audio

Berikut merupakan detail dari dimensi yang menjadi sumber data audio features spotify. Dimensi ini dapat ditemukan pada source db spotify.sqlite dengan keterangan atribut sebagai berikut.

DimAudio_Feature	Table Name	Feature	Table Type	Dimension	Display Name	Audio Feature	Database Schema	spotify.sqlite	Table Description	tabel audio features merupakan slowly changing dimension untuk berbagai informasi tentang fitur audio	Comment	berasal dari tabel audio_features di spotify.sqlite	Biz Filter Logic		Size	satu untuk setiap data audio feature	Generate Script?	Y
	Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Datatype	Size	Target Precision	Key?	FK To	NULL?	Source System	Source Schema	Source Table	Source Field Name	Source Datatype	
pk_audio_feature	pk_audio_feature	Kunci utama pengganti	-1	1, 2, 3...	key	int					PK ID	N	Derived	spotify	dbo	audio_features	id	int
id	id	Kunci bisnis dari sumber data untuk tabel audio_features (alias kunci alam)	-1	2jKoVU7VAwExU1Jh3w6P	key	int				N			spotify	dbo	audio_features	id	int	
acousticness	acousticness	fitur akustik dalam audio		180.000.007.152.557	1	int				N			spotify	dbo	audio_features	acousticness	int	
				https://api.spotify.com/v1/audio-analysis/2jKoVU7VAwExU1Jh3w6P	2	nvarchar	50			N			spotify	dbo	audio_features	analysis_url	nvarchar	
analysis_url	analysis_url	analisis link url untuk fitur audio		89.300.000.087.572	1	int				N			spotify	dbo	audio_features	analysis_url	int	
danceability	danceability	genre musik dance untuk fitur audio		219160	1	int				N			spotify	dbo	audio_features	danceability	int	
duration	duration	durasi atau lama waktu audio		513.999.999.599.489	1	int				N			spotify	dbo	audio_features	duration	int	
energy	energy	kekuatan audio fitur		0.0	1	nvarchar	50			N			spotify	dbo	audio_features	energy	int	
instrumentalness	instrumentalness	instrumental yang ada pada audio		11	1	int				N			spotify	dbo	audio_features	instrumentalness	nvarchar	
key	key	kunci fitur audio		595.999.999.383.894	1	int				N			spotify	dbo	audio_features	key	int	
liveness	liveness	pendeteksi hidup atau mati dari fitur audio		-507.999.992.370.906	1	nvarchar	50			N			spotify	dbo	audio_features	liveness	nvarchar	
loudness	loudness	kuat dari audio		1	1	int				N			spotify	dbo	audio_features	loudness	int	
mode	mode	mode yang digunakan dalam audio		282.999.992.370.905	1	int				N			spotify	dbo	audio_features	mode	int	
speechiness	speechiness	speechiness dari fitur audio		958.479.995.727.539	1	int				N			spotify	dbo	audio_features	speechiness	int	
tempo	tempo	tempo dari audio		4	1	int				N			spotify	dbo	audio_features	tempo	int	
time_signature	time_signature	waktu untuk menandakan beat audio		787.000.000.479.837	1	int				N			spotify	dbo	audio_features	time_signature	int	
valence	valence	valensi dalam fitur audio																

Gambar 4 Detailed Dimensional Modeling Dimensi Fitur Audio

3.2.4 Dimensi Genre

Berikut merupakan detail dari dimensi yang menjadi sumber data genre spotify. Dimensi ini dapat ditemukan pada source db spotify.sqlite dengan keterangan atribut sebagai berikut.

Table Name	DimGenre										
Table Type	Dimension	<div>Home Page</div>									
Display Name	Genre										
Database Schema	spotify.sqlite										
Table Description	tabel genre merupakan slowly changing dimension untuk berbagai informasi tentang genre										
Comment	berasal dari tabel genres di spotify.sqlite										
Biz Filter Logic											
Size	satu untuk setiap data genre										
Generate Script?	Y										
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Display Folder	ETL Rules	Comments	Datatype		
pk_genre	pk_genre	Kunci utama pengganti	-1	1, 2, 3...	key				int		
id	id	Kunci bisnis dari sumber data untuk tabel genres	-1	detroit his hon	1				text		

Gambar 5 Detailed Dimensional Modeling Dimensi Genre

3.2.5 Dimensi Track

Berikut merupakan detail dari dimensi yang menjadi sumber data track spotify. Dimensi ini dapat ditemukan pada source db spotify.sqlite dengan keterangan atribut sebagai berikut.

Table Name Dim Track		Dimension																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
----------------------	--	-----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

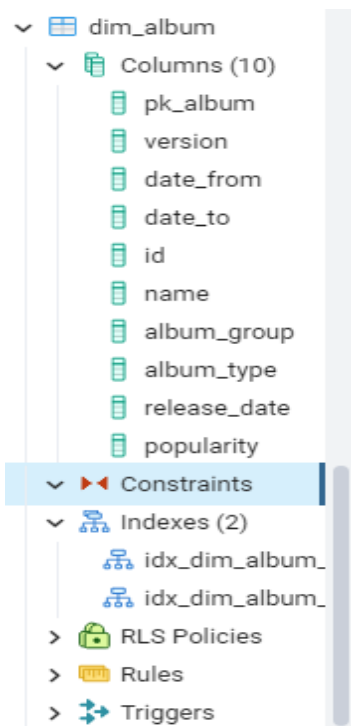
Gambar 6 Detailed Dimensional Modeling Dimensi Track

3.3. Dimensional Hierarchies

Pada bagian ini adalah menampilkan *dimensional hierarchi* pada lima dimensi, diantaranya adalah *albums*, *artist*, *audio_features*, *genres*, *tracks*.

3.3.1 Dimensi Album

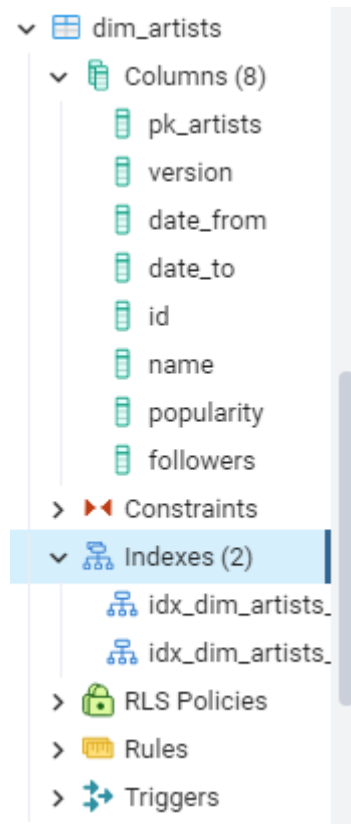
Berikut adalah tampilan dari dimensi albums dengan beberapa kolom data yaitu *pk_album*, *version*, *date_from*, *date_to*, *id*, *name*, *album_group*, *album_type*, *release_date* dan *popularity*.



Gambar 7 Dimensional Hierarchi Album

3.3.2 Dimensi Artis

Berikut adalah tampilan dari dimensi artist dengan beberapa kolom data yaitu *pk_artists*, *version*, *date_from*, *date_to*, *id*, *name*, *popularity* dan *followers*.



Gambar 8 Dimensional Hierarchi Artis

3.3.3 Dimensi Fitur Audio

3.3.4 Dimensi Genre

3.3.5 Dimensi Track

3.4. Dimensional Data Model

Dimensional data model merupakan metode pemodelan data menggunakan tabel fakta dan tabel dimensi yang digunakan dalam perancangan data untuk pengoptimalan dalam mengambil data pada *warehouse*. Pada proyek ini, model data dimensional yang digunakan adalah *star schema*. *Star schema* mempunyai tabel fakta yang terletak di bagian tengah dan

3.5. High-Level Source to Target Map

```
graph TD; Spotify[Spotify.sqlite] --> SCD_Pop1[SCD for popularity]; Spotify --> DataId[Data id tidak numerik]; Spotify --> SCD_Tempo[SCD for tempo]; Spotify --> SCD_Pop2[SCD for popularity]; Spotify --> SCD_Followers[SCD for followers]; Spotify --> SCD_Pop3[SCD for popularity]; SCD_Pop1 --> albums[albums]; DataId --> genres[genres]; SCD_Tempo --> SCD_Instrumental[SCD for instrumentalness]; SCD_Instrumental --> audio_features[audio_features]; SCD_Pop2 --> SCD_Followers; SCD_Followers --> artists[artists]; SCD_Pop3 --> tracks[tracks];
```

3.6. Detailed ETL Flow for Each Source to Target

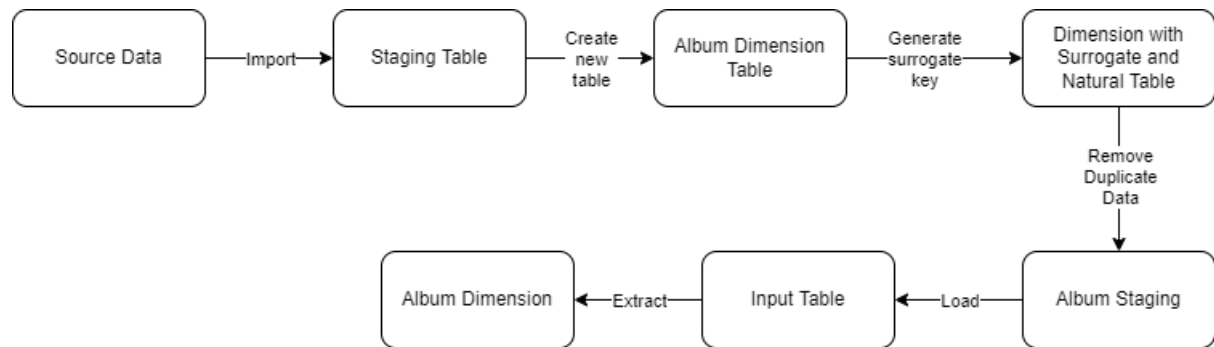
3.6.1 Dimensi Artis



- 9

3. *Generate surrogate key* untuk dimensi artis.
4. Hapus data-data yang mengalami duplikasi.
5. *Staging data table* yang sudah disimpan sebelumnya akan dimasukkan ke input tabel.
6. Setelah itu, data diekstrak ke dimensi artis.

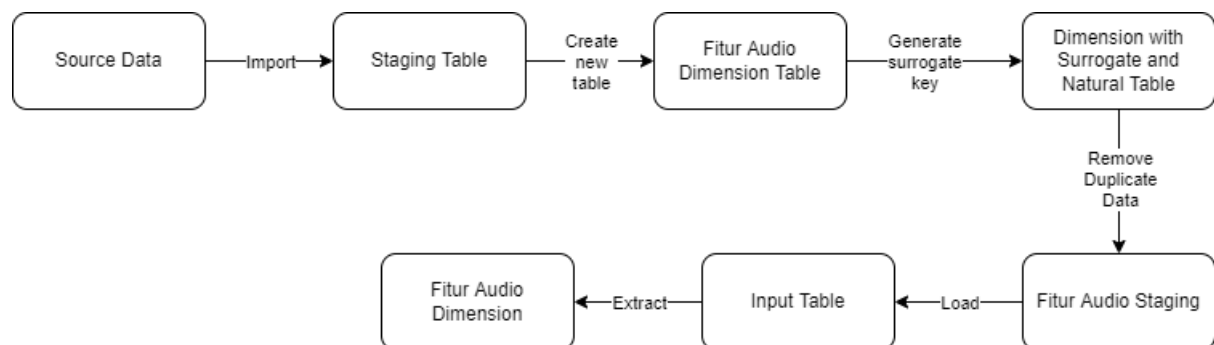
3.6.2 Dimensi Album



Gambar 11 Detailed ETL Flow Dimensi Album

1. Load source data ke PostgreSQL sebagai staging table
2. Buat tabel dimensi target, yaitu Album dimension, kemudian pilih atribut yang akan dimasukkan ke dalam tabel dimensi.
3. *Generate surrogate key* untuk dimensi album.
4. Hapus data-data yang mengalami duplikasi.
5. *Staging data table* yang sudah disimpan sebelumnya akan dimasukkan ke input tabel.
6. Setelah itu, data diekstrak ke dimensi album.

3.6.3 Dimensi Fitur Audio

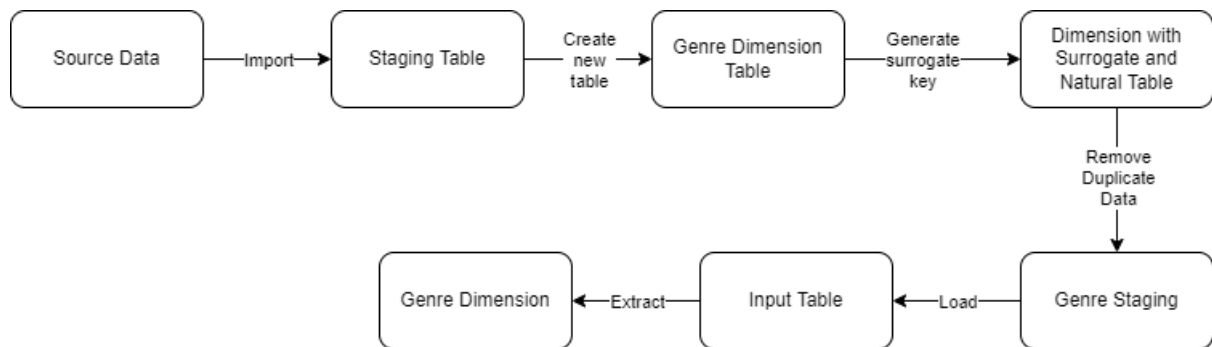


Gambar 12 Detailed ETL Flow Dimensi Fitur Audio

1. Load source data ke PostgreSQL sebagai staging table
2. Buat tabel dimensi target, yaitu Fitur Audio dimension, kemudian pilih atribut yang akan dimasukkan ke dalam tabel dimensi.
3. *Generate surrogate key* untuk dimensi fitur audio.

4. Hapus data-data yang mengalami duplikasi.
5. *Staging data table* yang sudah disimpan sebelumnya akan dimasukkan ke input tabel.
6. Setelah itu, data diekstrak ke dimensi fitur audio.

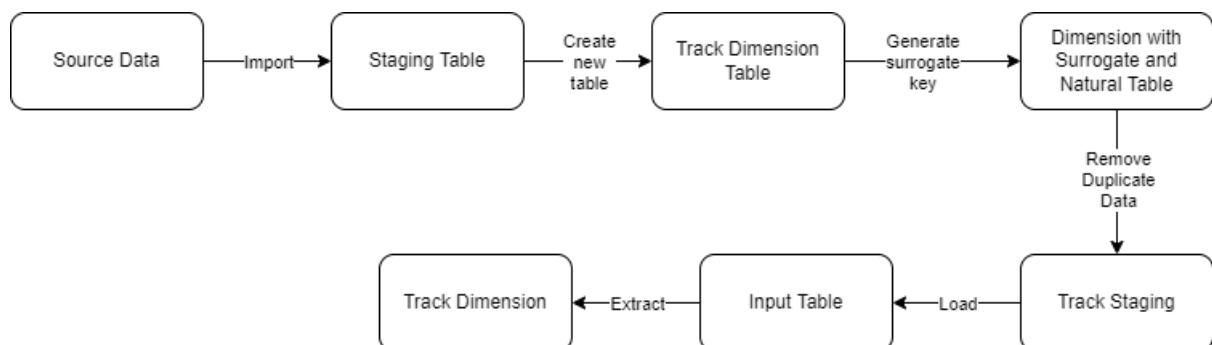
3.6.4 Dimensi Genre



Gambar 13 Detailed ETL Flow Dimensi Genre

1. Load source data ke PostgreSQL sebagai staging table
2. Buat tabel dimensi target, yaitu Genre dimension, kemudian pilih atribut yang akan dimasukkan ke dalam tabel dimensi.
3. *Generate surrogate key* untuk dimensi genre.
4. Hapus data-data yang mengalami duplikasi.
5. *Staging data table* yang sudah disimpan sebelumnya akan dimasukkan ke input tabel.
6. Setelah itu, data diekstrak ke dimensi genre.

3.6.5 Dimensi Track



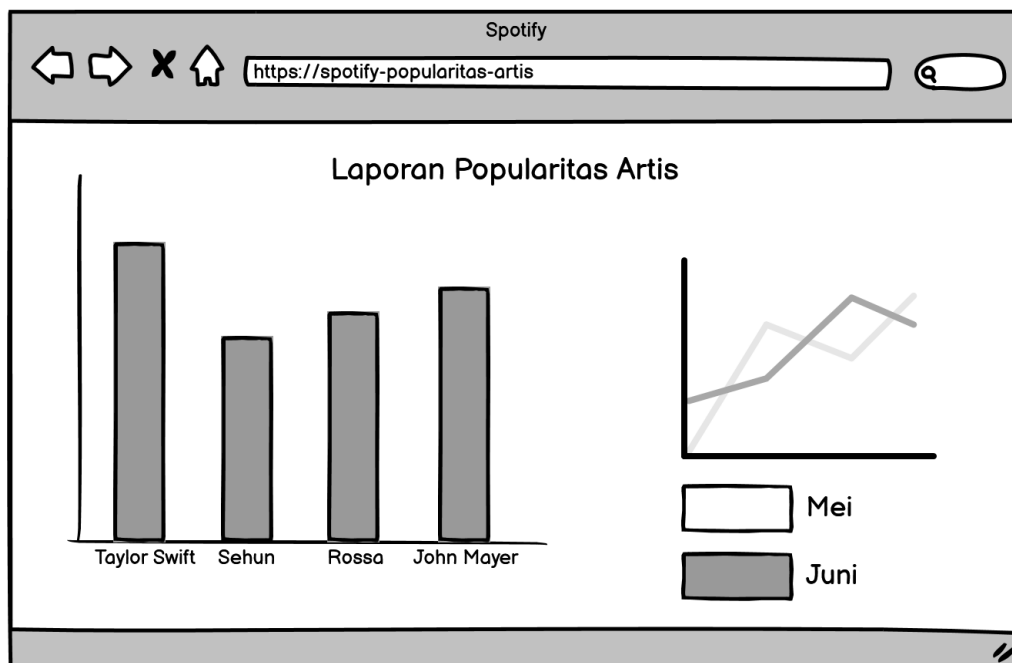
Gambar 14 Detailed ETL Flow Dimensi Track

1. Load source data ke PostgreSQL sebagai staging table
2. Buat tabel dimensi target, yaitu Track dimension, kemudian pilih atribut yang akan dimasukkan ke dalam tabel dimensi.
3. *Generate surrogate key* untuk dimensi track.
4. Hapus data-data yang mengalami duplikasi.

5. *Staging data table* yang sudah disimpan sebelumnya akan dimasukkan ke input tabel.
6. Setelah itu, data diekstrak ke dimensi track.

3.7. Business Intelligence Front End Mockup

Berikut merupakan mockup laporan kepopuleran artis perbulannya. Terdapat 2 chart pada mockup tersebut yaitu *bar chart* dan *line chart*. Pada *bar chart* ditampilkan total dari beberapa nama artis dengan perbandingan data kepopuleran artis yang berbeda-beda, sedangkan pada *line chart* ditampilkan grafik peningkatan kepopuleran artis pada bulan yang telah ditetapkan.



Gambar 15 Business Intelligence Front End Mockup

4. Implementasi

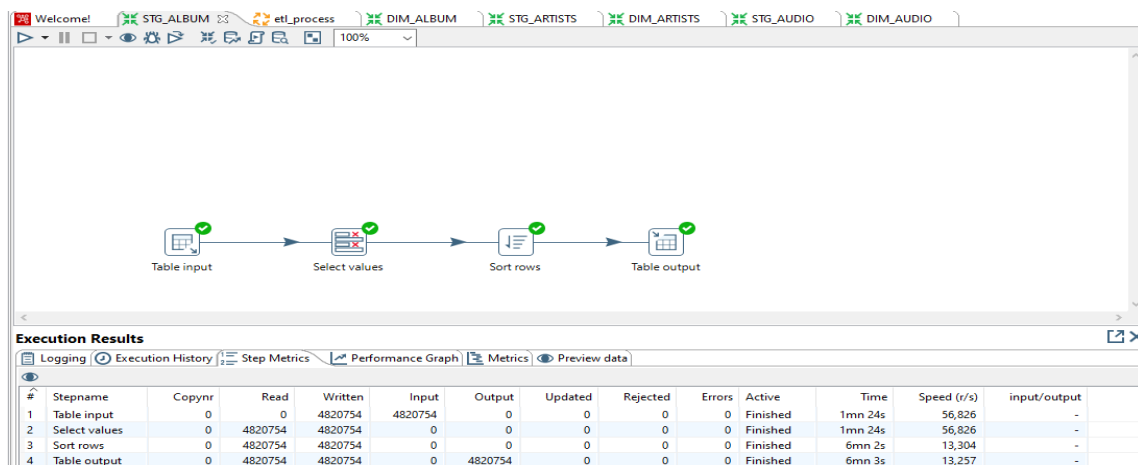
Bagian ini berisi ROLAP *schema*, kode program pada ETL, basis data MOLAP, dan *Business Intelligence Front End*.

4.1. ETL

Pada bagian ini adalah membuat proses ETL dari proses bisnis pelaporan artis terpopuler, dimana terdapat staging dan dimensi untuk setiap tabel.

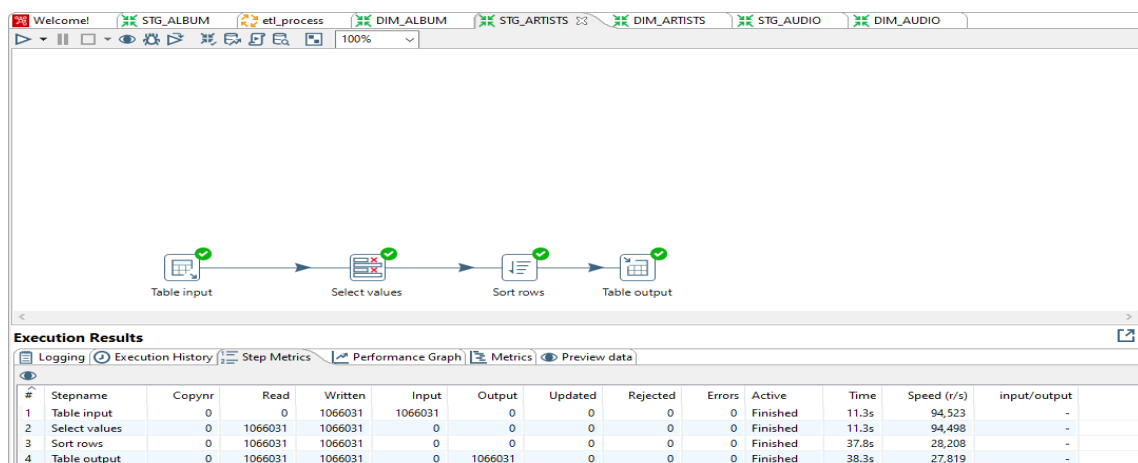
4.1.1 Staging

Pada bagian ini terdapat lima staging, yaitu stg_albums, stg_artists, stg_audio_features, stg_genres, dan stg_tracks.



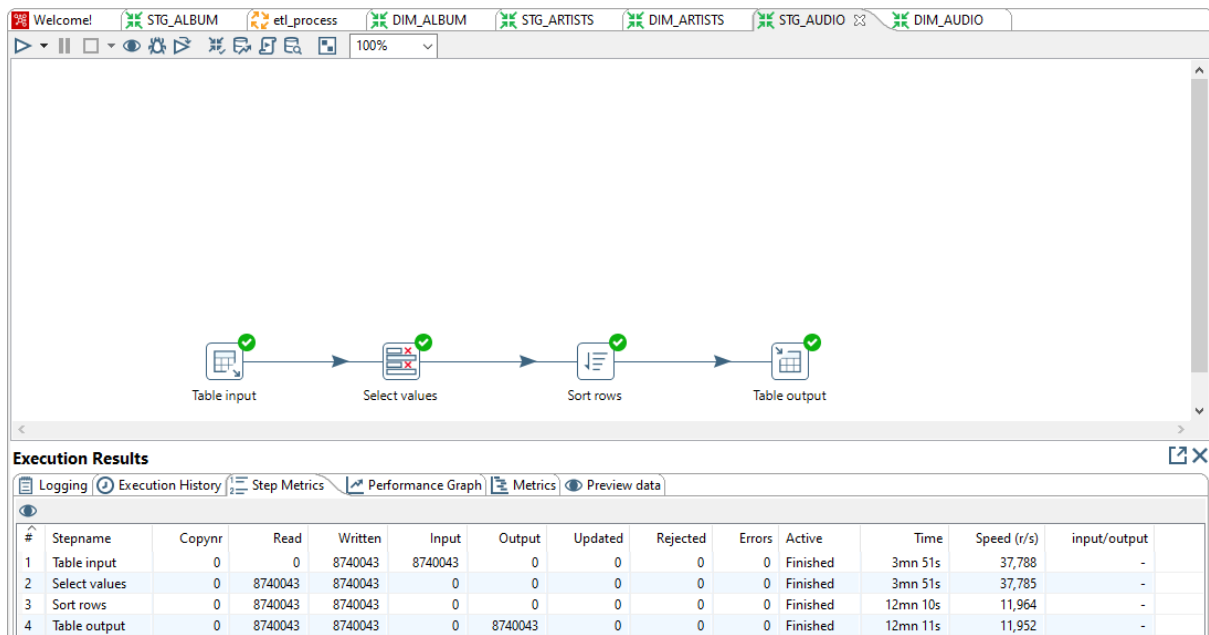
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	input/output
1	Table input	0	0	4820754	4820754	0	0	0	0	Finished	1mn 24s	56,826	-
2	Select values	0	4820754	4820754	0	0	0	0	0	Finished	1mn 24s	56,826	-
3	Sort rows	0	4820754	4820754	0	0	0	0	0	Finished	6mn 2s	13,304	-
4	Table output	0	4820754	4820754	0	4820754	0	0	0	Finished	6mn 3s	13,257	-

Gambar 16 Staging Album

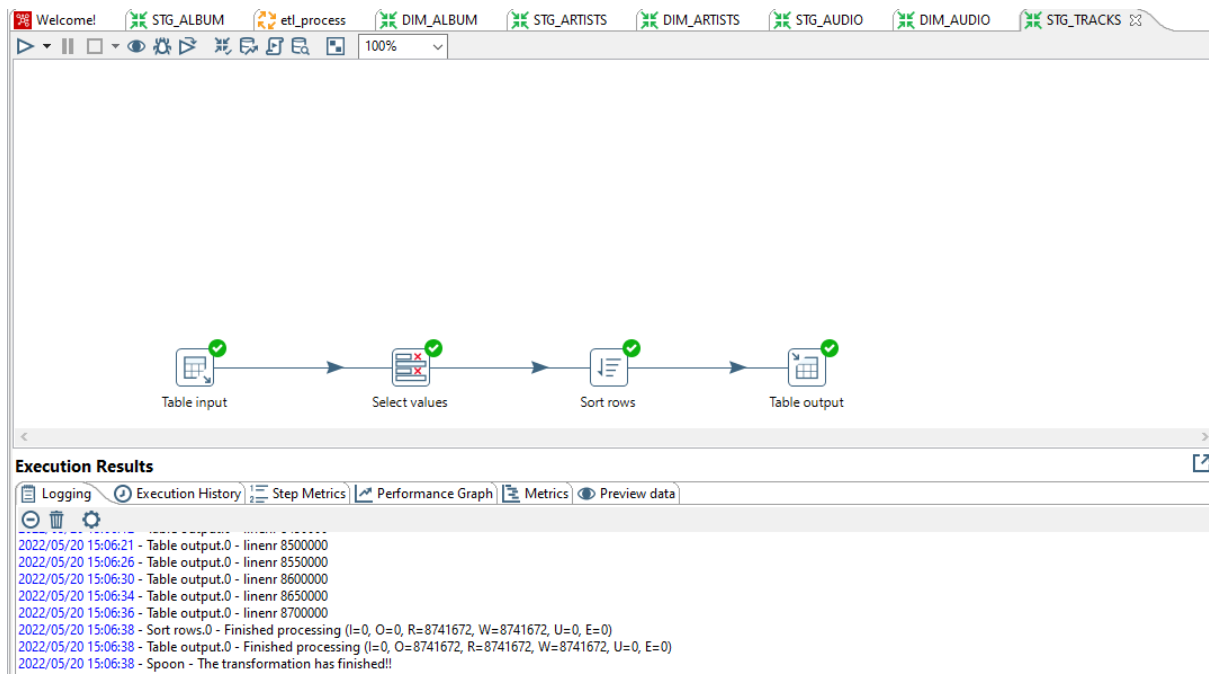


#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	input/output
1	Table input	0	0	1066031	1066031	0	0	0	0	Finished	11.3s	94,523	-
2	Select values	0	1066031	1066031	0	0	0	0	0	Finished	11.3s	94,498	-
3	Sort rows	0	1066031	1066031	0	0	0	0	0	Finished	37.8s	28,208	-
4	Table output	0	1066031	1066031	0	1066031	0	0	0	Finished	38.3s	27,819	-

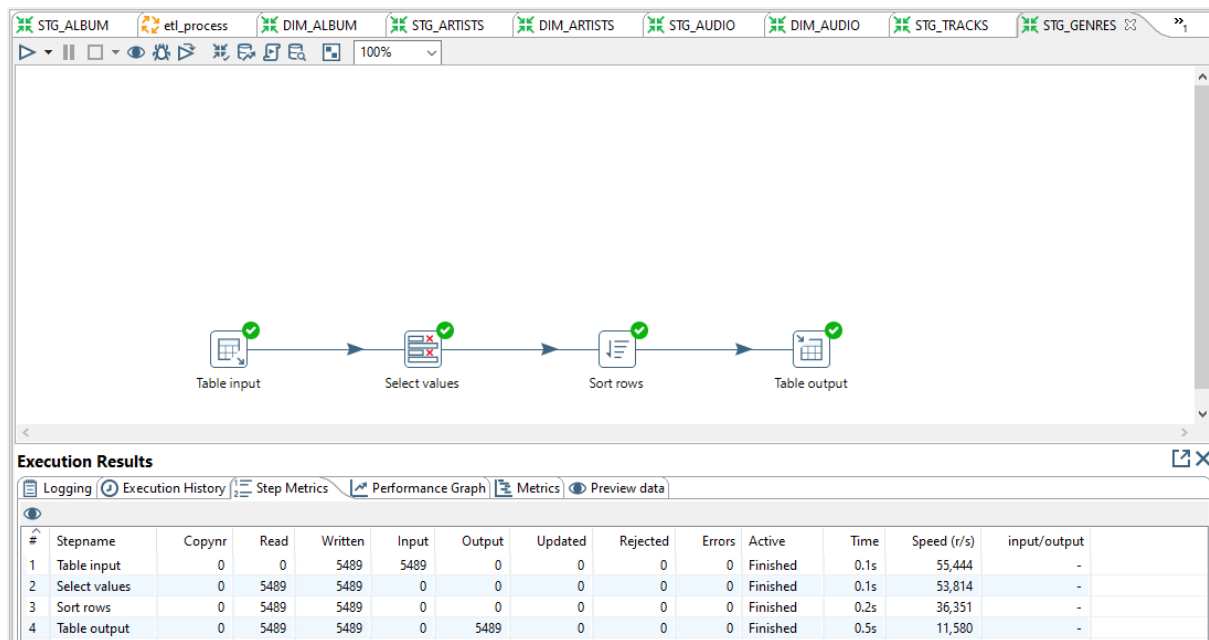
Gambar 17 Staging Artis



Gambar 18 Staging Fitur Audio



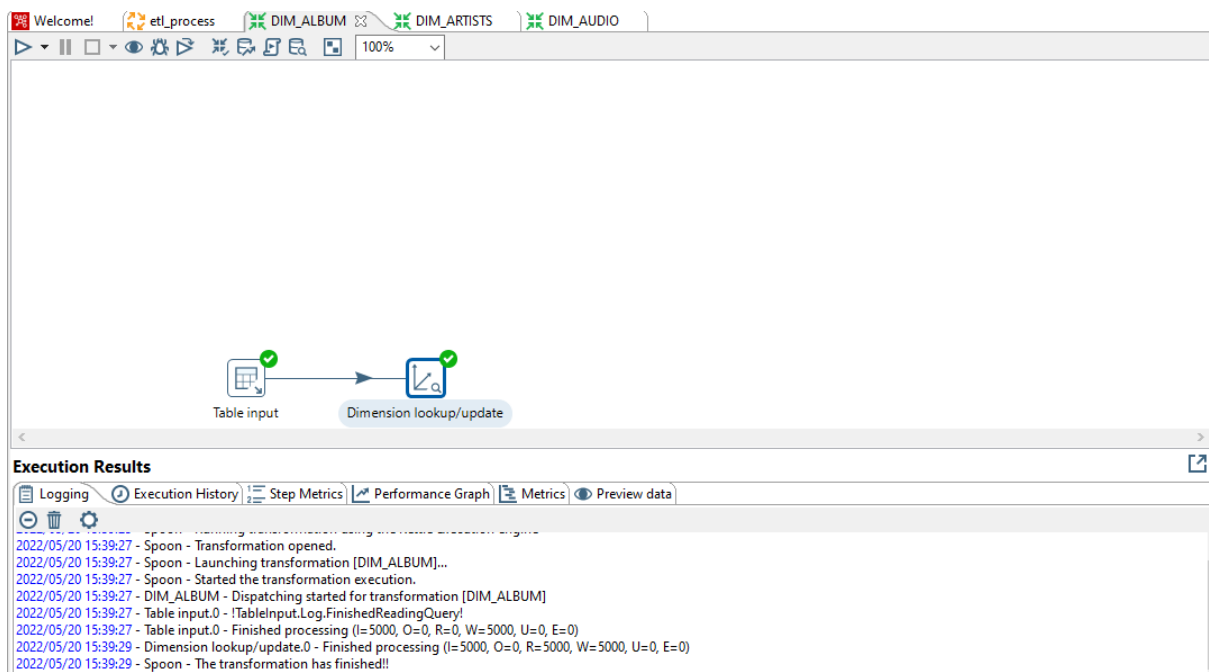
Gambar 19 Staging Track



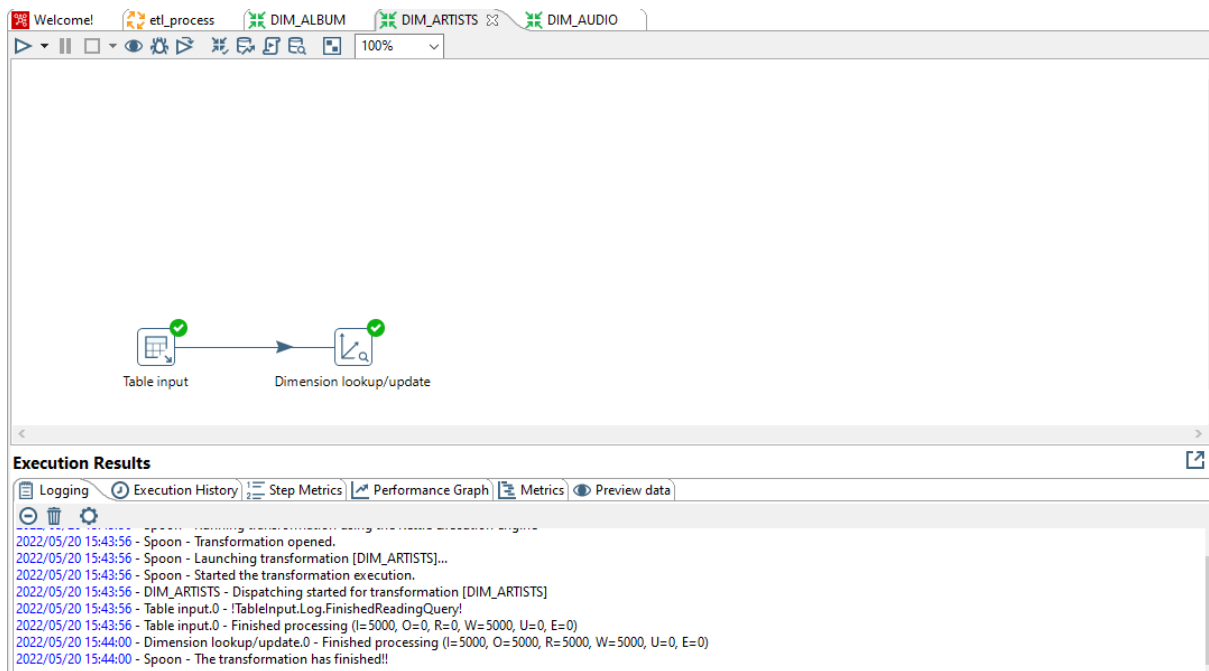
Gambar 20 Staging Genre

4.1.2 Dimension

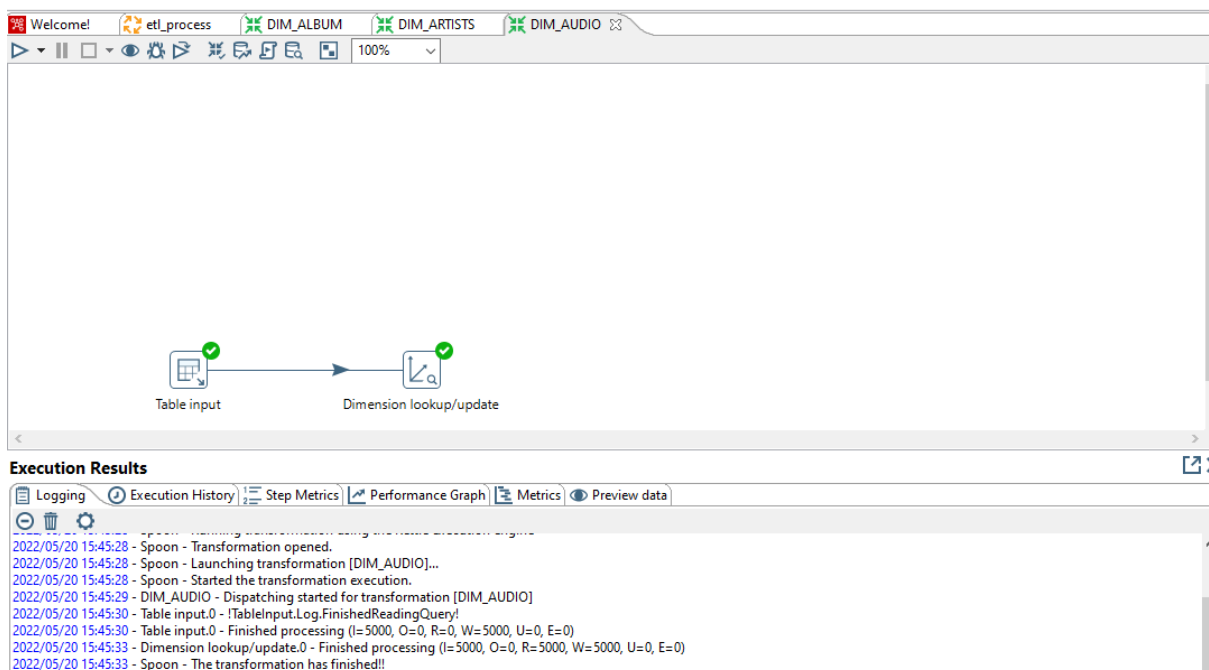
Pada bagian ini terdapat proses ETL dari setiap dimensi tabel, dimana terdapat dim_album, dim_artists, dim_audio_features, dim_genre, dan dim tracks.



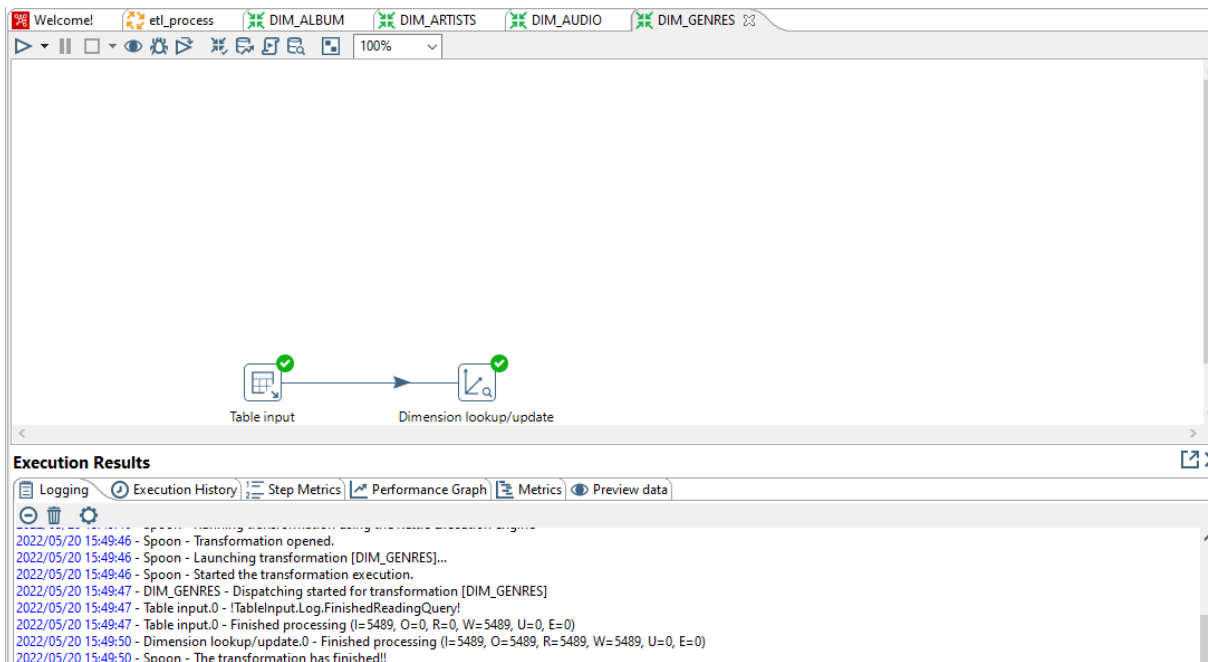
Gambar 21 Dimension Album



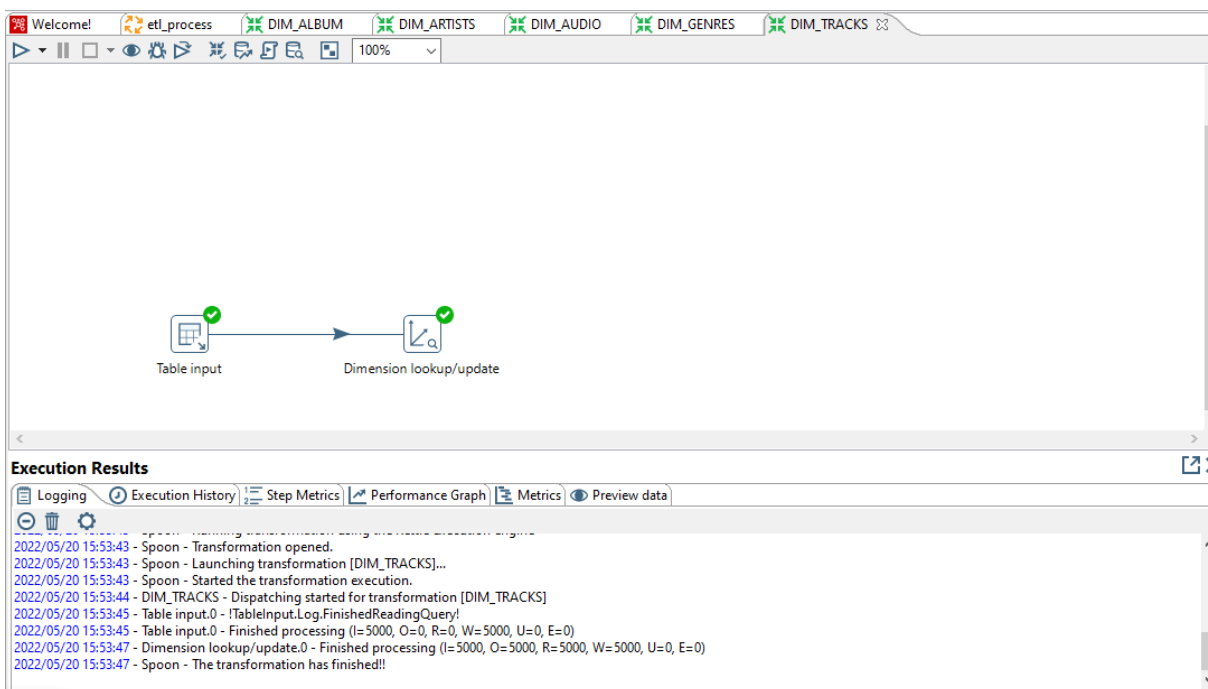
Gambar 22 Dimension Artis



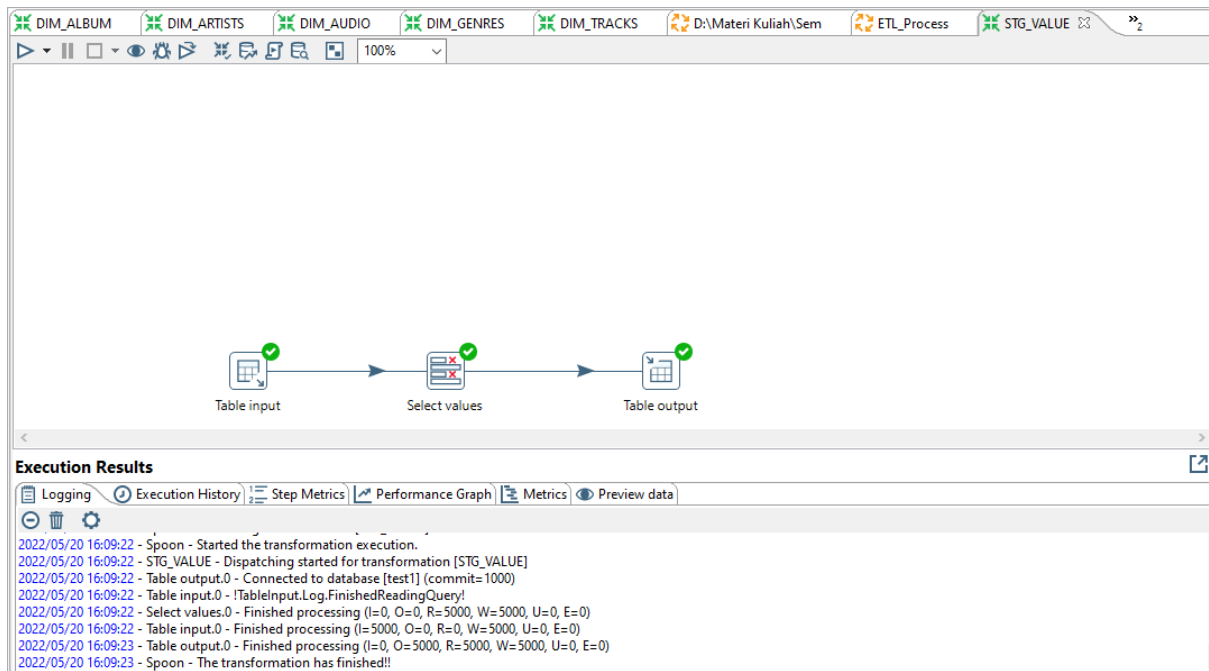
Gambar 23 Dimension Audio



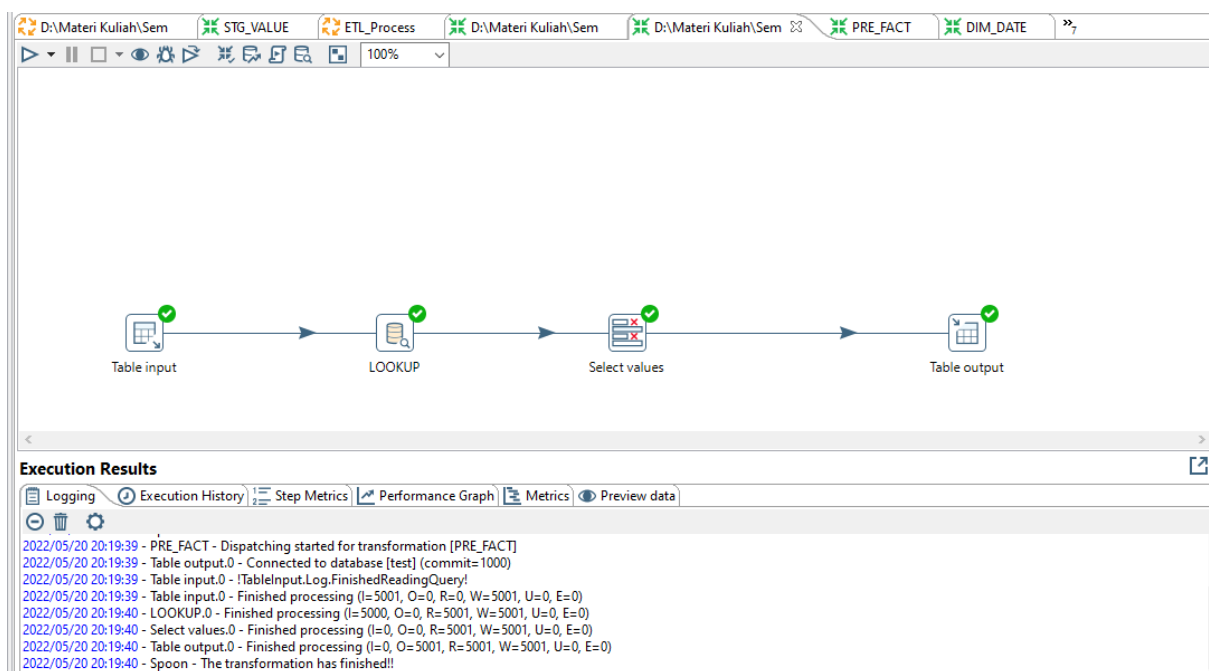
Gambar 24 Dimension Genre



Gambar 25 Dimension Track

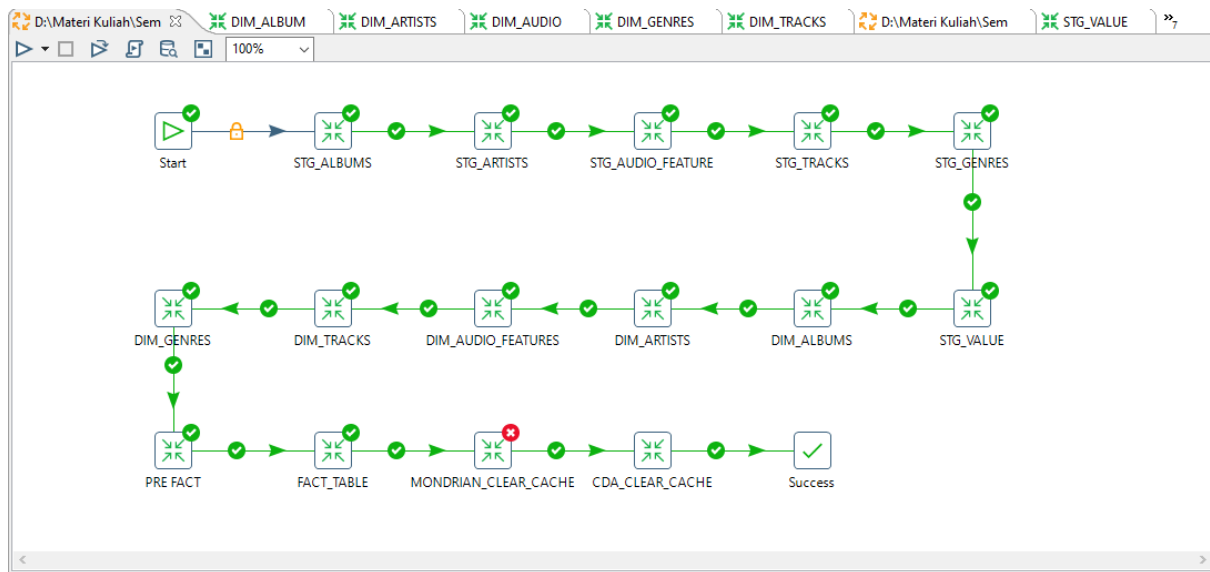


Gambar 26 Staging Value



Gambar 27 Pre Fact

Tampilan dibawah ini merupakan hasil akhir dari proses ETL.



Gambar 28 Proses ETL

4.2. Dashboard

5. Penutup

5.1. Kesimpulan

Sistem yang dibangun masih belum selesai dan belum dapat menyelesaikan masalah yang ada pada latar belakang.

5.2. Saran

Untuk penelitian selanjutnya, pengembang dapat menggunakan data lain seperti data warehouse selain menggunakan data spotify karena data spotify kurang mendukung dalam pembangunan data mart.

LAMPIRAN

<Opsional.>