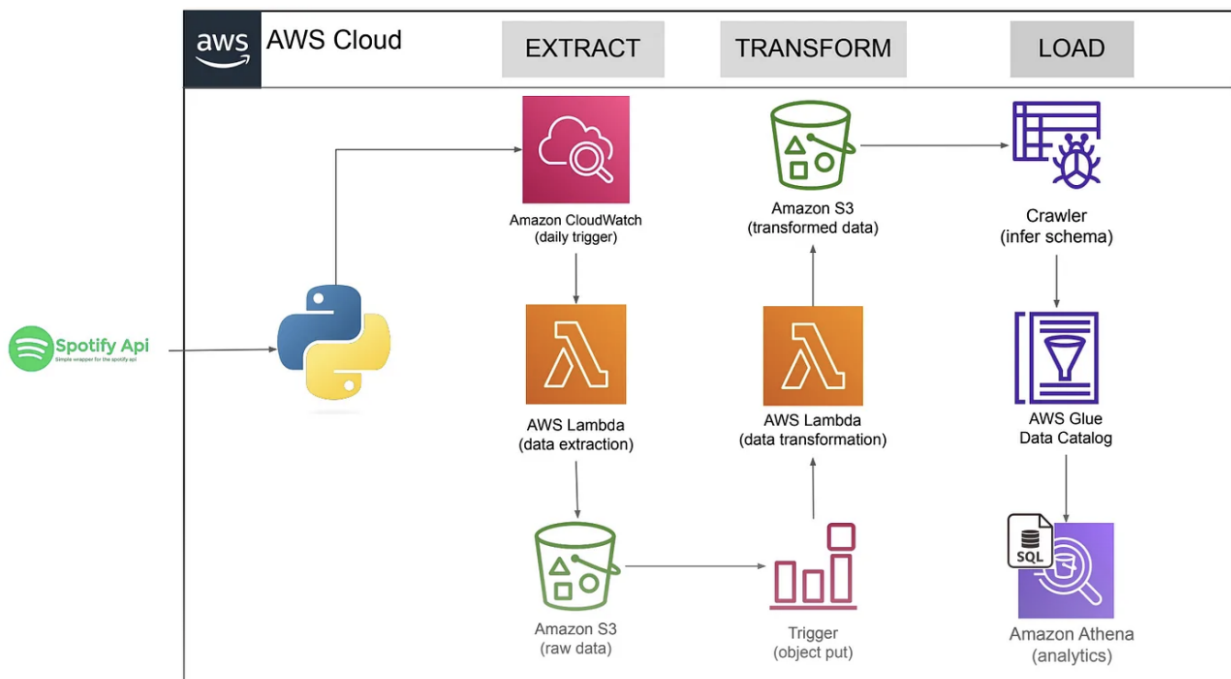


Spotify Data ETL Integration on AWS for Seamless Analytics

Details :

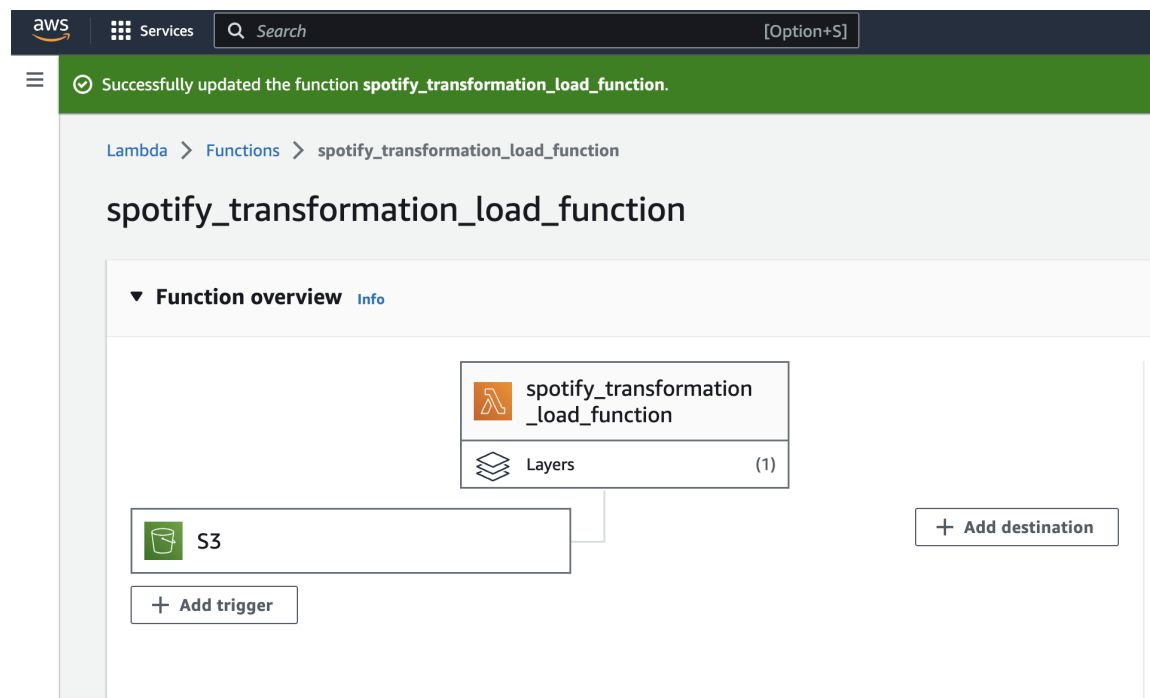
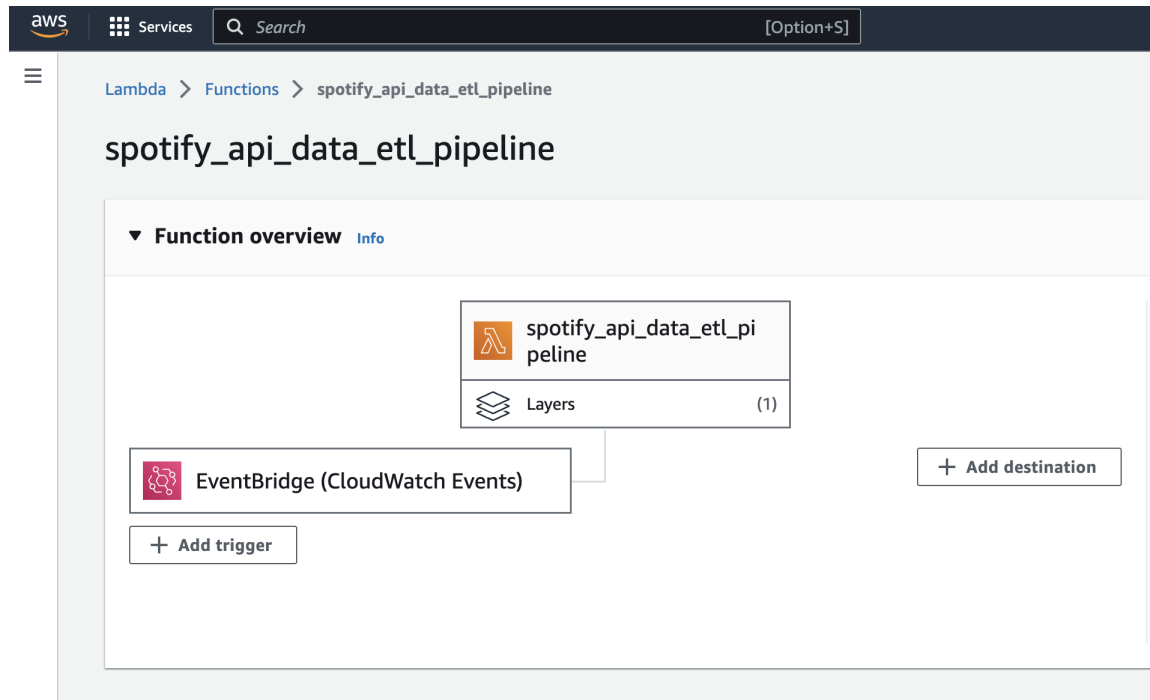
- Effectively integrated and utilized Spotify API to optimize data extraction
- Successfully deployed code onto AWS Lambda, establishing a smooth and automated data extraction process.
- Implemented triggers to facilitate automatic execution of the extraction process
- Developed robust transformation functions, enhancing data quality and utility
- Established automated triggers for seamless execution of transformation operations
- Effectively organized and stored data files on S3, ensuring optimal accessibility and management
- Skillfully constructed comprehensive analytics tables using Glue and Athena, facilitating insightful data analysis

Architecture :

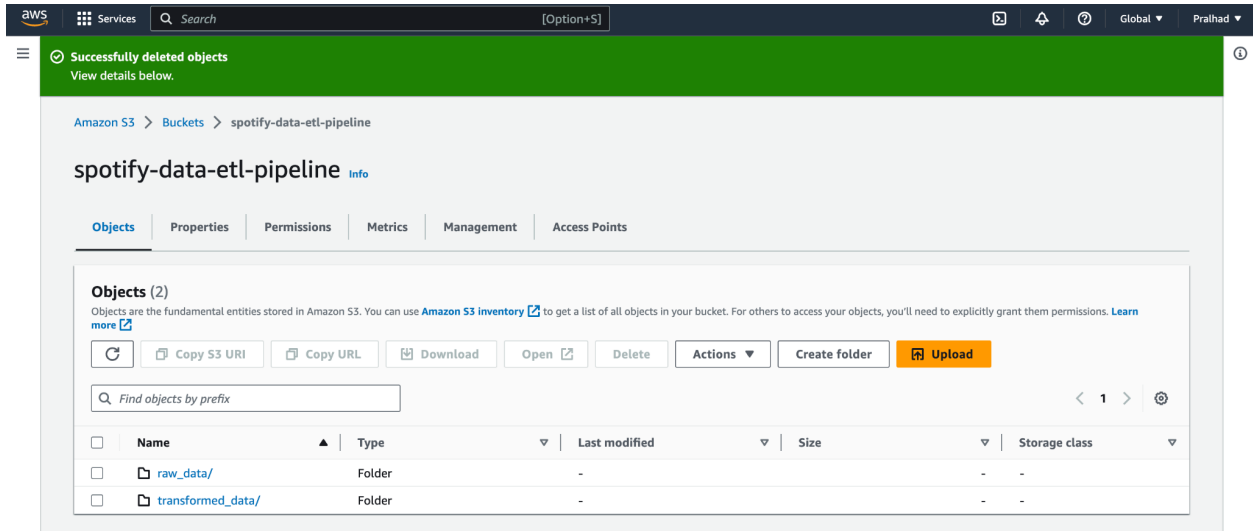


Output screenshots :

1. Event Triggers



2. AWS S3 Bucket Structure :



Amazon S3 > Buckets > spotify-data-etl-pipeline

spotify-data-etl-pipeline

[Info](#)

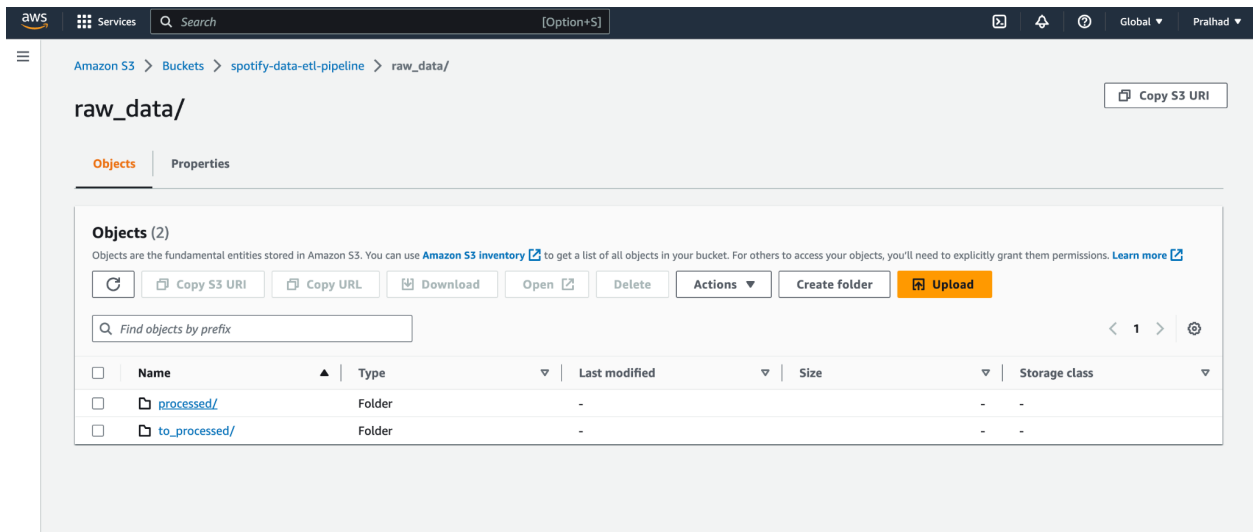
Objects | Properties | Permissions | Metrics | Management | Access Points

Objects (2)
Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

[Refresh](#) [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open](#) [Delete](#) [Actions](#) [Create folder](#) [Upload](#)

Find objects by prefix

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	raw_data/	Folder	-	-	-
<input type="checkbox"/>	transformed_data/	Folder	-	-	-



Amazon S3 > Buckets > spotify-data-etl-pipeline > raw_data/

[Copy S3 URI](#)

raw_data/

Objects | Properties

Objects (2)
Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

[Refresh](#) [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open](#) [Delete](#) [Actions](#) [Create folder](#) [Upload](#)

Find objects by prefix

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	processed/	Folder	-	-	-
<input type="checkbox"/>	to_processed/	Folder	-	-	-

aws

Services

Search

[Option+S]

Global

Pralhad

Amazon S3

Buckets

spotify-data-etl-pipeline

transformed_data/

transformed_data/

Copy S3 URI

Objects

Properties

Objects (3)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 Inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Refresh

Copy S3 URI

Copy URL

Download

Open

Delete

Actions

Create folder

Upload

Find objects by prefix

< 1 >

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	album_data/	Folder	-	-	-
<input type="checkbox"/>	artist_data/	Folder	-	-	-
<input type="checkbox"/>	songs_data/	Folder	-	-	-

aws

Services

Search

[Option+S]

Global

Pralhad

Amazon S3

Buckets

spotify-data-etl-pipeline

transformed_data/

songs_data/

songs_data/

Copy S3 URI

Objects

Properties

Objects (4)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 Inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Refresh

Copy S3 URI

Copy URL

Download

Open

Delete

Actions

Create folder

Upload

Find objects by prefix

< 1 >

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	songs_transformed_2023-08-15 06:46:04.010595.csv	csv	August 14, 2023, 23:46:05 (UTC-07:00)	8.7 KB	Standard
<input type="checkbox"/>	songs_transformed_2023-08-15 07:16:52.747261.csv	csv	August 15, 2023, 00:16:53 (UTC-07:00)	8.7 KB	Standard
<input type="checkbox"/>	songs_transformed_2023-08-16 17:45:04.436171.csv	csv	August 16, 2023, 10:45:05 (UTC-07:00)	8.7 KB	Standard
<input type="checkbox"/>	songs_transformed_2023-08-16 18:09:06.756594.csv	csv	August 16, 2023, 11:09:08 (UTC-07:00)	8.7 KB	Standard

3. AWS Glue data log :

The screenshot shows the AWS Glue console interface. On the left, there's a navigation menu with options like 'Getting started', 'ETL jobs', 'Visual ETL', 'Notebooks', 'Job run monitoring', 'Data Catalog tables', 'Data connections', 'Workflows (orchestration)', 'Data Catalog', 'Databases', 'Tables', 'Stream schema registries', 'Schemas', 'Connections', 'Crawlers', 'Classifiers', 'Catalog settings', and 'Data Integration and ETL'. The main area displays the 'spotify_db' database properties, including its name, description, location, and creation date (August 15, 2023 at 06:52:29 UTC). Below this, there's a section for 'Tables (3)' with a filter bar and a table listing the tables: 'album_data', 'artist_data', and 'songs_data'. Each table entry shows its name, database, location, classification, deprecated status, and links to view data and data quality.

Name	Database	Location	Classification	Deprecated	View data	Data quality
album_data	spotify_db	s3://spotify-data-etl-	CSV	-	Table data	View data quality
artist_data	spotify_db	s3://spotify-data-etl-	CSV	-	Table data	View data quality
songs_data	spotify_db	s3://spotify-data-etl-	CSV	-	Table data	View data quality

4. AWS Athena - SQL querying :

The screenshot shows the AWS Athena console interface. On the left, there's a navigation menu with options like 'Data', 'Data source', 'Database', 'Tables and views', 'Tables (3)', 'Views (0)', 'Query results', and 'Query stats'. The main area displays a SQL query: 'SELECT * FROM "spotify_db"."songs_data" limit 10;'. The query has been executed successfully, and the results are shown in a table with 10 rows. The table columns are '#', 'song_id', 'song_name', 'duration_ms', and 'url'. The results show the first 10 songs from the 'songs_data' table, including 'Seven (feat. Latto) (Explicit Ver.)', 'Cruel Summer', 'What Was I Made For? [From The Motion Picture "Barbie"]', 'LALA', 'Dance The Night (From Barbie The Album)', 'vampire', and 'Bad Religion (feat. Aesop Rock) [From Barbie The Album]'.

#	song_id	song_name	duration_ms	url
1	7x9aauaA9cu6tyfpHnqDLo	Seven (feat. Latto) (Explicit Ver.)	184400	https://open...
2	1BxfuPKGuaTgP7aM0Bbdwr	Cruel Summer	178426	https://open...
3	6wf7YU7cx8SPRlWeSeK0Q	"What Was I Made For? [From The Motion Picture "Barbie"]"	222369	https://open...
4	7ABLnD53cQK00mhcaOUVG	LALA	197920	https://open...
5	1vYXt7V5jH9JIM5oRRo7vA	Dance The Night (From Barbie The Album)	176579	https://open...
6	3k79jB4aGmMDUQzEwa46Rz	vampire	219724	https://open...
7	74118B73u7T0aFvuaQ...	Bad Religion (feat. Aesop Rock) [From Barbie The Album]	202350	https://open...