

	Total de Usuários	% Masculino	% Feminino	Dias Úteis	Média de Reproduções	Total de Streamings	Maior Resultado (Més)	Maior Resultado (semana)	Idade Média Usuários
	962K	70,0%	30,0%	262	2,58	16,3M	Dez - 2020	Terça-Feira	34,03
Overview Metadados	Gênero Rock Electronic Pop Rap Metal Country Folk RnB Reggae Jazz New Age Punk	8.808.2 1.975.595 1.484.501 891.687 846.604 697.086 389.822 363.087 191.751	31 Streamings 1.385.919	Série Histórica de S Diferença (mês vs mê 1.337.368 4% 3% 4% abr 2020 jul 2020	5treamings	Ano 2008 2009 2007 2010 2010 2005 2004 2001 2003 2012	Reproduções 1.866.767 1.831.380 1.521.470 1.404.941 1.117.292 1.051.640 856.557 668.005 663.105 598.341 568.789 433.967 424.133 339.347 322.746	Artista Kings of Leon Justin Bieber Coldplay Modest Mouse Metric Lady Gaga Radiohead Jack Johnson Kid Cudi Eminem Daft Punk John Mayer The New Pomographers The Smiths	Reproduções 570.134 185.071 170.581 167.042 157.351 155.558 154.942 124.004 120.653 106.865 105.715 105.176 102.212
	●World ●Blues			1.117,292			322.439 319.925	Rise Against	97.772 92.432
	E	170.658		319.925 1.739 51.772	856.557 663.705 598.341 424.133 778	± 1997 ± 1995	253.641 179.654	Metallica Katherine Jenkins	91.752 91.448
	7	74.884				± 1994	173.938	Widefille	90.563
	5	091	21.113 11.			± 1998 ± 1993	173.778 143.971	Muse Guns N' Roses	83.721 82.687
0	3	17.240	19	80 2		202C Total	16.344.469	Total	16.344,469

Metadados								
Zone	Table	Column	Туре	Description				
Trusted	music_info	acousticness	Double	Self Explainable				
Trusted	music_info	artist	String	Artist Name				
Trusted	music_info	danceability	Double	Parameter Of A Song				
Trusted	music_info	duration_ms	Int	Duration Of Song In Miliseconds				
Trusted	music_info	energy	Double	Parameter Of A Song				
Trusted	music_info	genre	String	Music Genre				
Trusted	music_info	instrumentalness	Double	The Higher Less Vocals				
Trusted	music_info	key	Int	Key Song Was Made				
Trusted	music_info	liveness	Double	Higher Means Live Music				
Trusted	music_info	loudness	Double	Higher Means Louder				
Trusted	music_info	mode	Int	Song Parameter				
Trusted	music_info	name	String					
Trusted	music_info	speechiness	Double	Higher Means More Vocals				
Trusted	music_info	spotify_id	String	Unique Value				
Trusted	music_info	spotify_preview_url	String	Access To Song Demo				
Trusted	music_info	tags	String	Genres Sep By Comma				
Trusted	music_info	tempo	Double	Bps Higher Means Faster Beats Per Second				
Trusted	music_info	time_signature	Int	Song Parameter				
Trusted	music_info	track_id	String	Unique Value				
Trusted	music_info	ts_proc	Bigint					
Trusted	music_info	ts_proc_partition	Bigint					
Trusted	music_info	valence	Double	Song Parameter				
Trusted	music_info	year	Int	Year Release				
Trusted	user_info	ts_proc	Bigint					
Trusted	user_info	ts_proc_partition	Bigint					
Trusted	user_info	user_id	String	Unique Id User				
Trusted	user_listening_history	playcount	Bigint	Register Play				

Pipeline para ingestão de dados e processamento (batch)

- 1. Dados são obtidos do ambiente on-premise e sincronizados para o S3 via AWS Cli.
- 2. A "Transient Zone" é a primeira camada de ingestão de dados, onde ocorre a catalogação das origens e tipos de dados. Essa é uma zona temporária e os arquivos são apagados após o carregamento na zona de "Raw Data".
- 3. O EventBridge Scheduler está configurado para iniciar o pipeline.
- 4. Na zona de "Raw Data", os dados são armazenados de forma "bruta", sem qualquer modificação.
- 5. O AWS Step Functions é responsável pela orquestração de jobs.
- 6. Uma lambda function é destinada para o catalogamento dos arquivos que chegam ao bucket S3 (Transient Zone).
- 7. Em seguida, outra lambda é acionada. Iniciando um cluster EMR, posteriormente, jobs são adicionados para o processamento dos arquivos.
- 8. Novas tabelas são geradas e armazenadas no formato Parquet no bucket (S3) Trusted Zone.
- Os dados presentes nessa área já passaram por processamentos e transformações necessárias para serem utilizados, assegurando sua qualidade e tornando-os precisos e confiáveis.
- 10. Consultas SQL realizadas pelo Athena.
- 11. Dados são carregados para o Power BI, para a criação de dashboards.
- 12. Dashboards e relatórios são acessados por outras áreas/clientes.