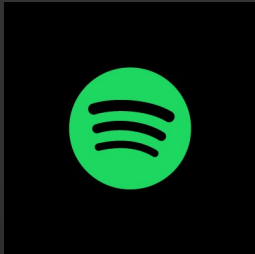


- Home
- Search
- Library
- New playlist
- Liked songs

- PYSPARK
- DATABRICKS
- BIGDATA
- FINAL PROJECT
- STREAMING
- SONGS



Data 603

Spotify Charts Analysis



Custom sorting

TITLE

ALBUM

ADDED ON

Data 603 Final Project
Submitted By: Karan Ajay Pisay & Srashti Soni

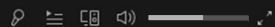


DATA 603 – FINAL PROJECT
Artist



1:44

2:29



Introduction:

- Spotify, which began in Sweden, has grown to become one of the world's most popular music streaming services.
- Since 2006, the company has grown in popularity, and in the first quarter of 2022, it has 182 million paying users.
- The enormous number of Spotify users, made it possible to build a big content database and extract useful patterns from Spotify charts and track metadata.
- The analysis conducted on the Spotify datasets provides significant insights into the data and popularity of the tracks, artist and it also helps comprehend the types of tracks that were popular in which region across the globe.
- Scope of this project is to provide information to the music companies and artists to understand their audience well and produce songs according to the likability of the audience.

Data Gathering, Data cleaning & Data Wrangling

- For the current project we have used 3 datasets from Kaggle. The dataset meets the project requirement as it sums up to 3.7 GB after combining them together.
- We have used Pyspark to perform extract and transform operations. To gather the data, we had to chunk the large CSV file into small chunks and then load it into data bricks. Once all eleven chunks are loaded into the data bricks, we extracted all of them into 3 data frames.
- The extracted data is then cleaned and used for further analysis.
- After an overview of the data, we observed that the data had to be cleaned and managed according to the schemas. Moreover, we had to ensure that the null values were eliminated and the data was clean and ready for analysis without any garbage and duplicate values.

Overview of the data:

- As we can see the data ranges from the following dates. It contains the list of songs, artists, metadata of the music, region of the songs, and several more attributes.

	begin	end
0	2017-01-01	2021-12-31

	begin	end
0	1922	2021

- This is a general overview of the artists that appeared on the top200 charts of Spotify.

	artist	chart
0	Marama	top200
1	The Chainsmokers	top200
2	Gilda	top200
3	Shelley FKA DRAM	top200
4	Ed Sheeran	top200
...
36143	Spelloppmaker	top200
36144	Kwiat Jabloni	top200
36145	RAPK	top200
36146	D Law	top200
36147	Lil AK	top200

Artists with highest number of appearance on the charts:

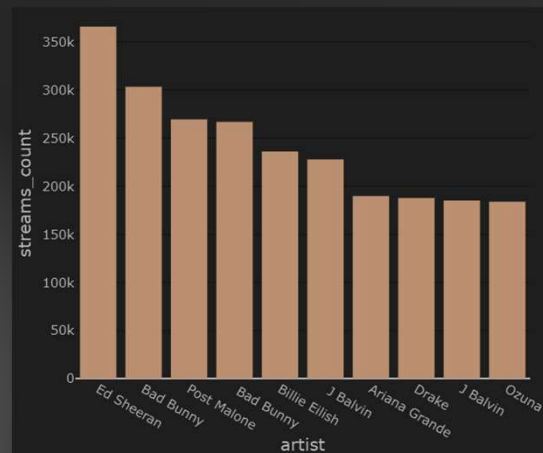
- As we can see in the table, Ed Sheeran appeared on the charts highest number of times followed by Bad Bunny and Post Malone. And after some further research, we were able to find out that Ed Sheeran, Bad Bunny and Post Malone were on the chart's highest times because of the following songs:

Ed Sheeran : Shape Of You, Perfect, Thinking Out Loud

Bad Bunny : Drivers License with(Olivia Rodrigo)(Spanish)

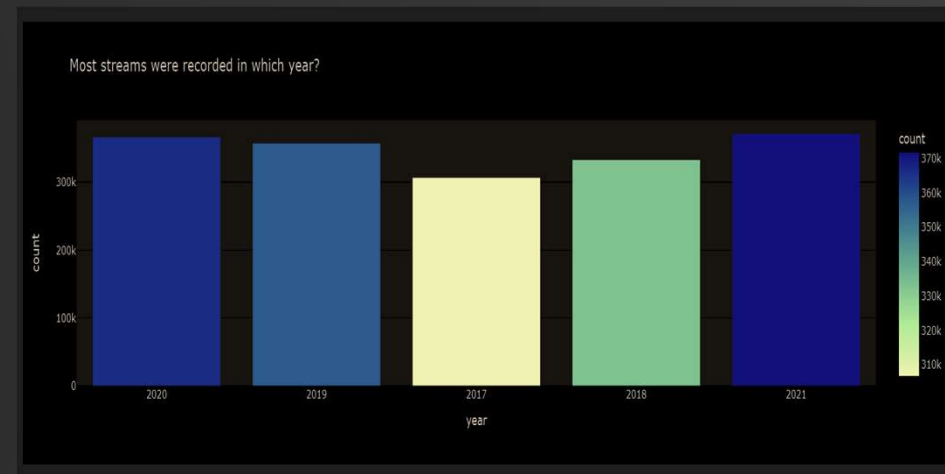
Post Malone: Rockstar and SunFlower.

	artist ▲	streams_count ▲
1	Ed Sheeran	366151
2	Bad Bunny	303549
3	Post Malone	269649

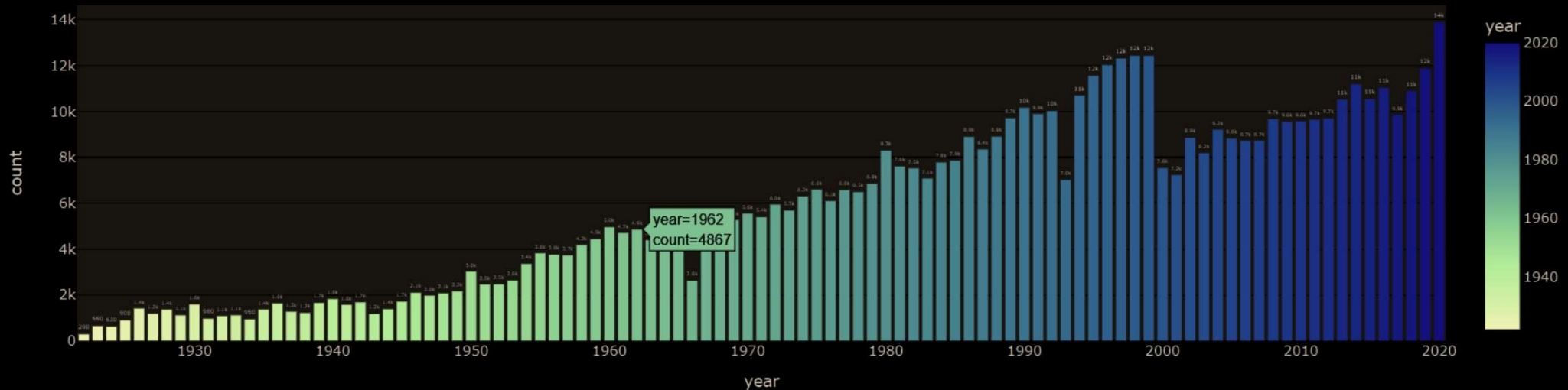


Most streams were recorded in which year?

As we can see that the maximum streams were recorded in the year 2021. However, we cannot completely ignore that the year 2020 and 2021 were the years of COVID pandemic and study suggests that people were more active on social media during this time and in the past couple of years, Spotify also introduced new offers and plans to promote and increase their customer base. Hence there has been a higher number of activity since 2019 in the Spotify dataset.



Tracks released throughout the years



- A general overview of the data set shows that the maximum songs were released in the year 2020.
- We believe that this significant impact on the number of songs released has happened due to the pandemic.
- As people were quarantined, artist had better accessibility to people as a lot of people introduced themselves to their creative sides.

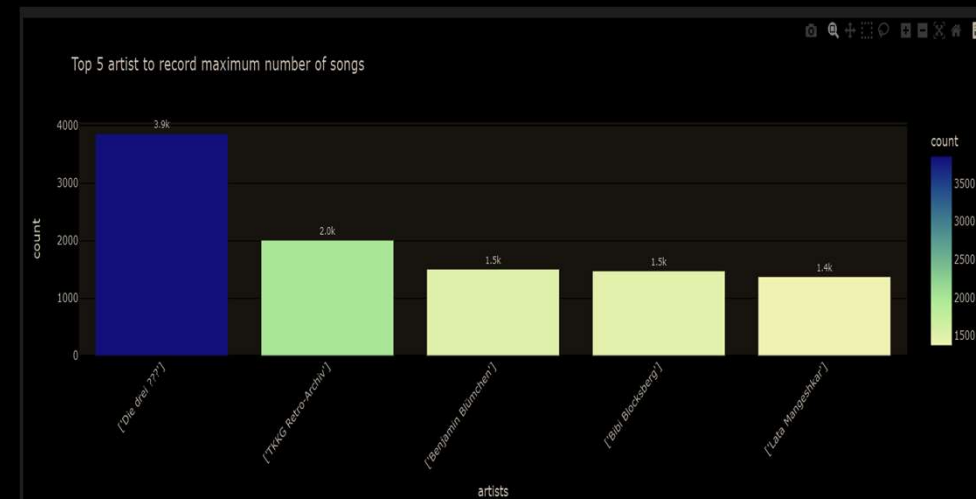
Artist with highest number of streams:

- Post Malone, Ed Sheeran and Drake have also manages to acquire the highest number of streams on Spotify. Further research showed that Post Malone Is Just The Second Musician To See Five Songs Reach One Billion Streams On Spotify after Justin Bieber.

artist	Top_streams
Post Malone	27068948579
Ed Sheeran	23635228040
Drake	20816590900
Billie Eilish	17598907282
Bad Bunny	16804299322
Ariana Grande	16033413364
J Balvin	15013042021
The Weeknd	14236255102
Bad Bunny	13529695586
XXXTENTACION	12656361852
Dua Lipa	12637009722
Juice WRLD	12453514331
J Balvin	11333909788
Travis Scott	10083054470
Olivia Rodrigo	9517707837
Ozuna	9466881998
Justin Bieber	9344338313

Top 5 artist to record maximum number of songs

FUN Fact: Lata Mangeshkar has lived 94 yrs. that is 33000 days approx. but after some google research it was found that she has over 50,000 sounds recorded under her name.



Artist with maximum number of tracks to feature on charts

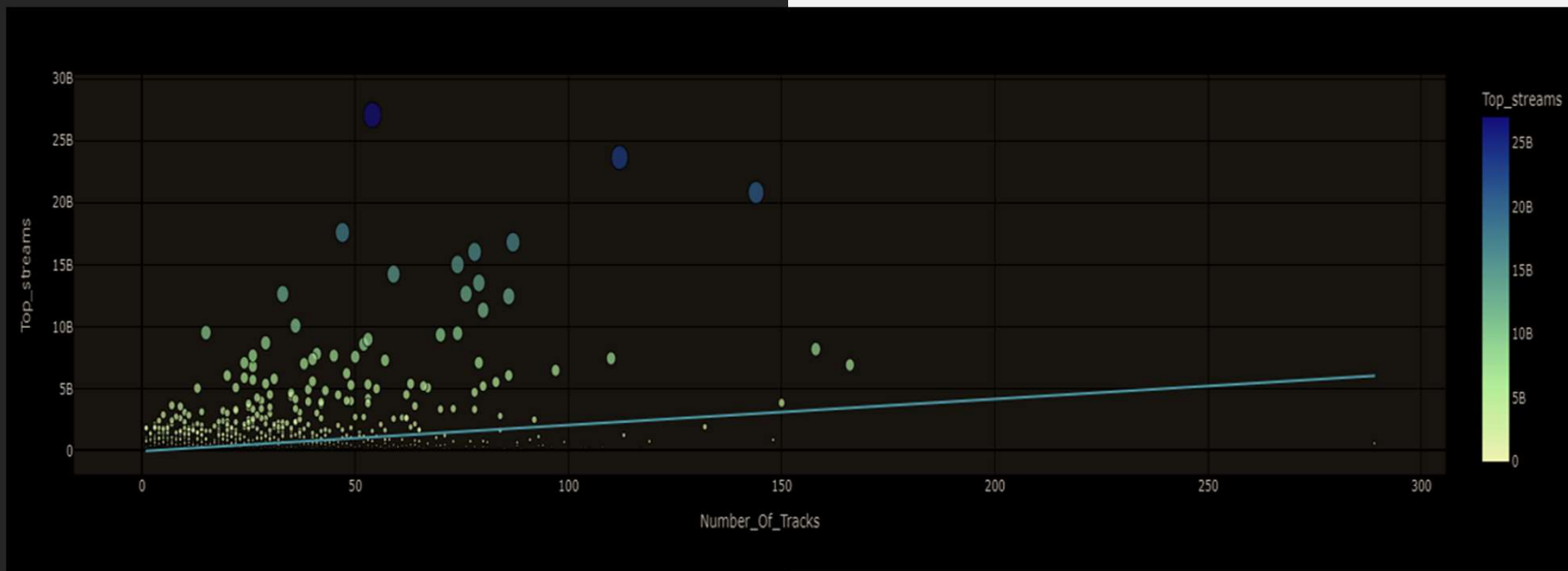
Jul and BTS, it's no surprise that they're so popular. They make music that appeals to a wide range of people, regardless of their interests, and they promote the notion that it's vital to support and love oneself. Furthermore, topics like self-love was key in the past couple of years.

artist	Number_Of_Tracks
Jul	289
BTS	166
Taylor Swift	158
Future	150
Ufo361	148
Drake	144
Die drei ???	138
Capital Bra	132
TWICE	119
Sevn Alias	117

Artist with highest number of charts in a day:

As we can see in the above table, Linkin Park came across 850+ charts on 07/21/2017. And after a short google search, we found out that in case of Linkin Park, XXXTENTACION and Avicii; they are appeared on the maximum number of charts on Spotify right after the day they died.

date	artist	count
2017-07-21	Linkin Park	1263
2021-08-30	Kanye West	1263
2018-06-19	XXXTENTACION	1238
2018-06-29	Drake	1180
2021-08-29	Kanye West	1156
2018-06-30	Drake	1104
2021-09-03	Drake	1095
2021-08-31	Kanye West	1091
2017-03-05	Ed Sheeran	1081
2017-03-04	Ed Sheeran	1079
2017-03-06	Ed Sheeran	1075
2017-03-03	Ed Sheeran	1048
2017-03-07	Ed Sheeran	1022
2018-07-01	Drake	1001
2017-03-08	Ed Sheeran	985
2017-03-09	Ed Sheeran	958
2018-07-02	Drake	953
2021-09-04	Drake	951



Finding the relation graph
between the number of
songs and the number of
streams:

- From the above graph we can evidently see that some streams have managed to get high number of streams or fall in the category of 15 to 20 million plus even though they have no composed more than 5 tracks and on other hand there is a artist who has composed nearly 300 songs but not managed to get the streams. This shows that there is not much correlation between the number of songs and the streams.

Highest number of streams for a single track on a single day and about 'Olivia Rodrigo'

In the top 20, we see many of them are Christmas songs played a day prior and post 25th DEC. However, All the other 11 are by Olivia Rodrigo.

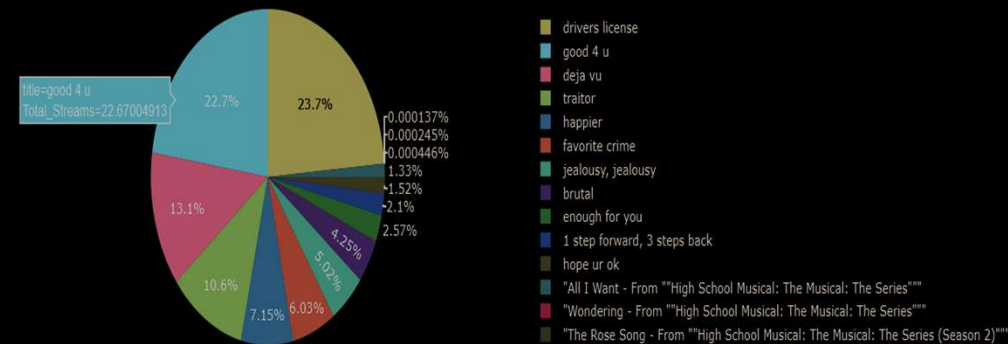
title rank	artist	date	streams
Easy On Me 1	Adele	2021-10-15	19749704
All I Want for Ch... 1	Mariah Carey	2020-12-24	17223237
Last Christmas 2	Wham!	2020-12-24	15813799
drivers license 1	Olivia Rodrigo	2021-01-15	13714177
drivers license 1	Olivia Rodrigo	2021-01-14	13701177
drivers license 1	Olivia Rodrigo	2021-01-13	13628143
All I Want for Ch... 1	Mariah Carey	2020-12-25	13573228
drivers license 1	Olivia Rodrigo	2021-01-12	13412192
Easy On Me 1	Adele	2021-10-16	13271124
good 4 u 1	Olivia Rodrigo	2021-05-25	12586645
good 4 u 1	Olivia Rodrigo	2021-05-26	12428245
Last Christmas 2	Wham!	2020-12-25	12415012
Girls Want Girls ... 1	Drake	2021-09-03	12384750
good 4 u 1	Olivia Rodrigo	2021-05-27	12292053
Santa Tell Me 3	Ariana Grande	2020-12-24	12229331
good 4 u 1	Olivia Rodrigo	2021-05-24	12200817
good 4 u 1	Olivia Rodrigo	2021-05-21	12192526
drivers license 1	Olivia Rodrigo	2021-01-16	12110121

Olivia Rodrigo was trending in 2021 because she had released her 2 biggest hits "GOOD 4 U" and "DRIVERS LICENSE" in 2021. She first appeared on top200 charts on 27th Dec 2019.

	year	streams
1	2021	13714177
2	2021	13701177
3	2021	13628143

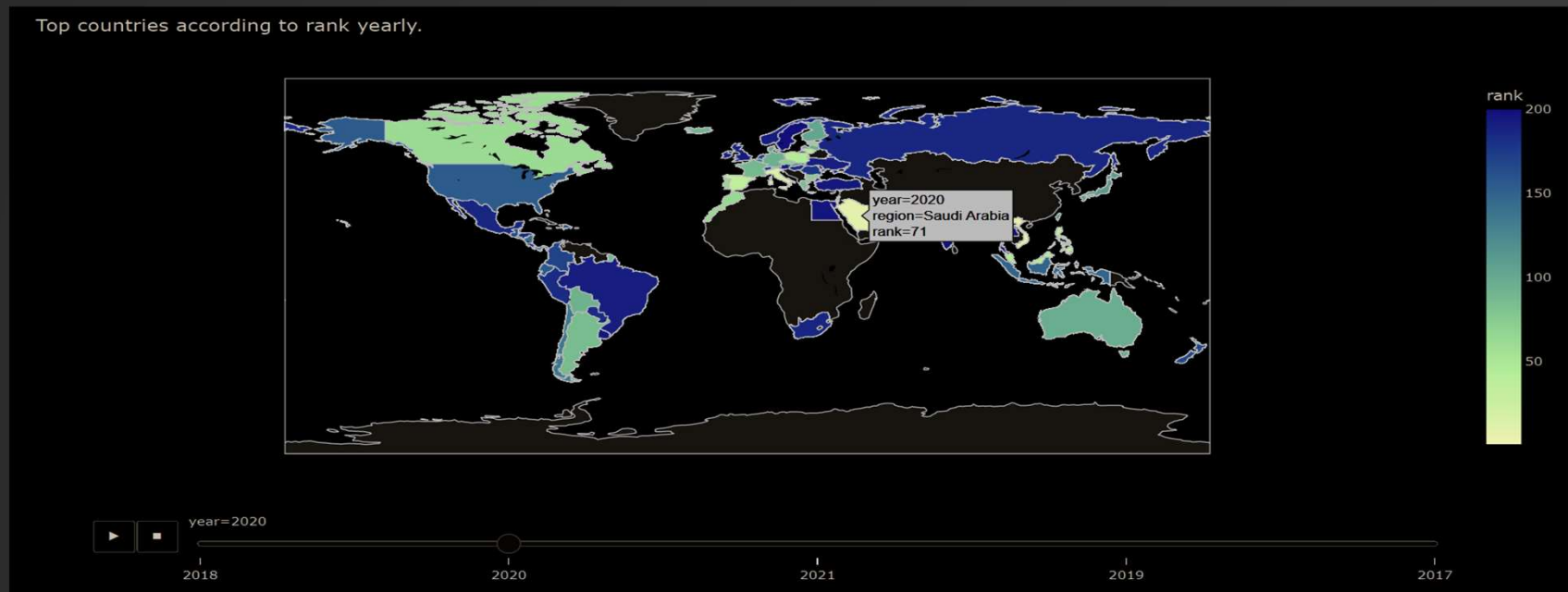
Out of all her streams maximum streams came from the following two songs:

Distribution of streams Olivia Rodrigo



Which region has appeared on top200 charts rank wise:

After playing the visualization, we have observed that European countries and North American and South American countries have been consistent on the Top 200 charts in the year 2017-2021. The songs produced in these contraries have been ranked in the top200 through the years.



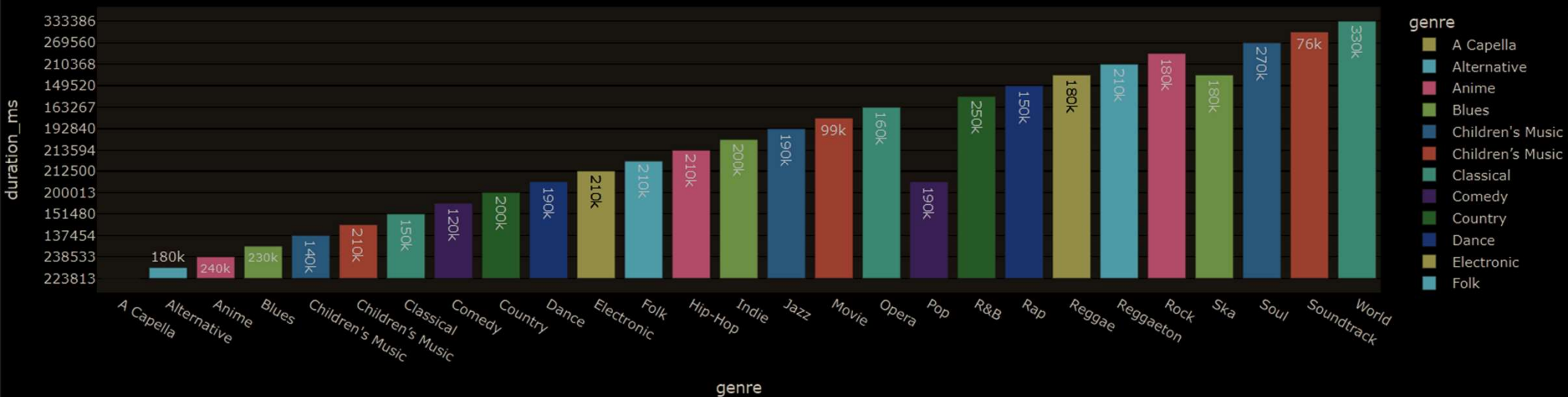
Which Genre has the maximum duration of music produced over the years?

- Here we can see that maximum duration of sound produced is from the world genre followed by soundtrack.
- World Genre is western popular music clubbed with traditional music.
- Soundtrack is a genre. It belongs to shared tradition or a set of conventions.

	genre	count
0	Comedy	9681
1	Soundtrack	9646
2	Indie	9543
3	Jazz	9441
4	Pop	9386
5	Electronic	9377
6	Children's Music	9353
7	Folk	9299
8	Hip-Hop	9295
9	Rock	9272
10	Alternative	9263
11	Classical	9256
12	Rap	9232

Approximate count of genre according to songs released over the years.

Duration according to Genre

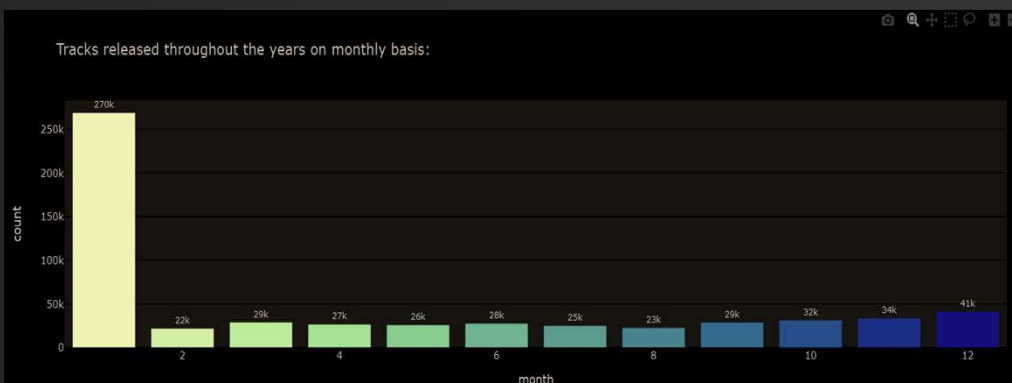


Most popular Genres amongst all!



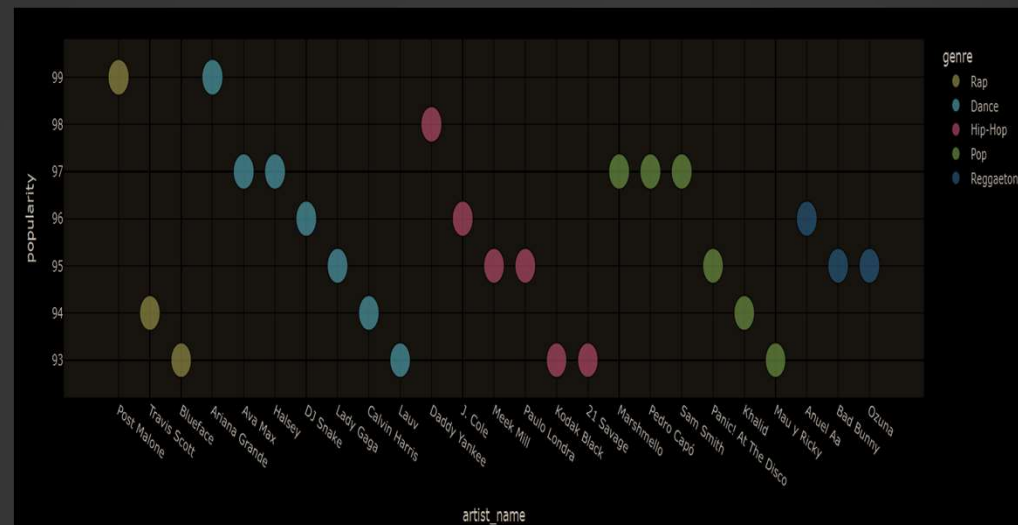
When are the maximum songs released throughout the year and why?

- January, because in the beginning of the year, everyone is also getting back into the swing of things and eager to find the next big thing. Furthermore, initial days of the year, the minds of your listeners are open to new things.



Artists with different genre and their popularity

- As we can see that different artist are popular for different styles of music. Below graph shows that how various artist excelled in their Genre and climbed their way to the top.

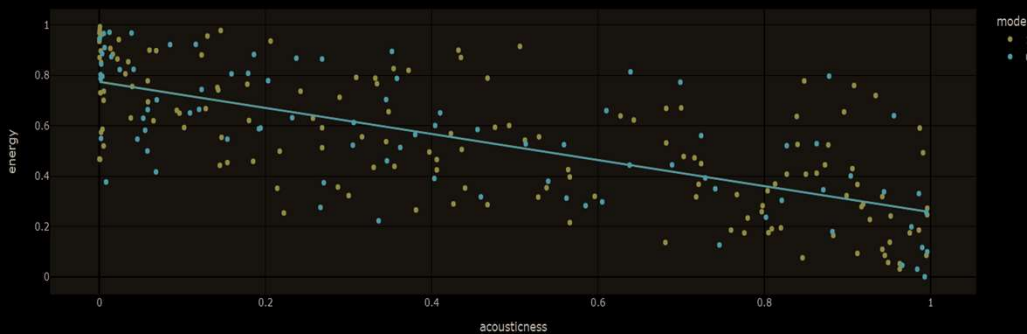


Now let's delve into the minute details of the music:

Energy VS acoustiness of the music:

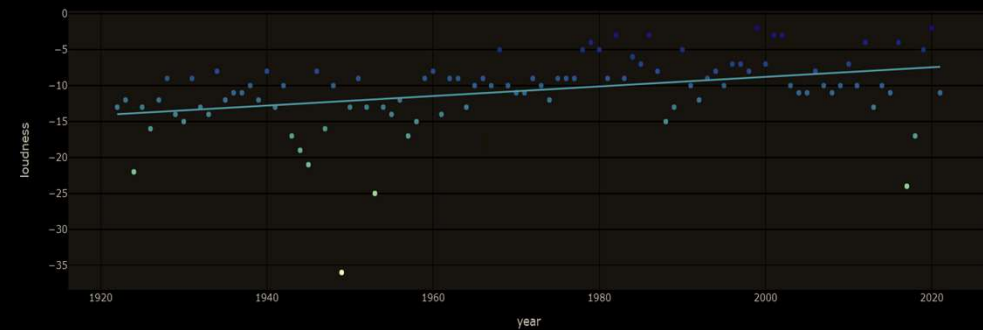
- Here we observe that the energy of the songs is inversely proportional to the acoustiness as the energy of the song drops the acoustiness increases.

	energy	acoustiness	mode
477900	0.639	0.6270	1
179998	0.854	0.0347	1
269259	0.443	0.1450	1
575792	0.289	0.9190	1
489664	0.744	0.1230	0

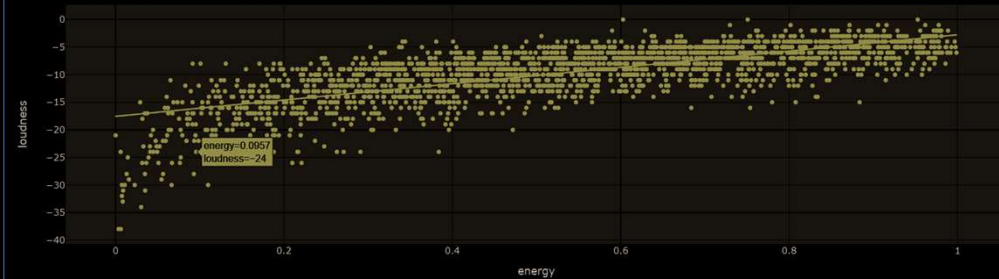


Loudness of the songs over the years:

- As the data suggests, the loudness of the songs over the years have increased but there is no significant raise in the loudness.



Loudness and Energy:



A background image showing a smartphone with the Spotify logo on its screen. A pair of white earbuds is resting on the phone. A semi-transparent circle is overlaid on the left side of the image, containing text and a list of facts.

Some Fun Facts About Spotify:

- Spotify has 182 million premium users globally in the first quarter of 2022, up from 158 million in the same quarter of 2021. Spotify's subscription base has grown significantly in recent years, more than doubling since early 2017.
- For the second year in a row, Bad Bunny was the most-streamed artist in the world on Spotify at the end of 2021.
- Bad Bunny became the all-time most-streamed artist on Spotify in a single day on Friday, May 6th, according to Spotify. Within 24 hours, he had 183 million Spotify streams.
- "Despacito" was the song that stayed in first position for the longest period of time. This song has been streamed over 1.9 billion times on Spotify.

Conclusion:

- After analyzing the data, we were able to conclude that there are humungous number to deal with and gather information from the music industry. Music industry being one of the oldest and the most popular sector contains so many unknown and hidden facts.
- Every artist has to out stand of the crowd to ensure success and they must deliver the right type of music depending on their audience to ensure they remain on the charts throughout their time.
- Some artist produce music which promotes self-love while the other find their ways by composing music based on the occasions and festivals.
- Producing music consistently, does not ensure success or higher number of streams. Popularity of the song is simply based on the music and the audience's likability.
- The majority of popular songs are more danceable and energetic. Such tunes are appropriate for lightening our mood or for playing at celebrations.
- Ed Sheeran , Bad Bunny and post Malone are considered as the most consistent and highly successful artist of the time.
- Out of the millions of artists around the globe, only 2097 were able to make in on the top200 charts.
- Despacito was on the top for the maximum number of days
- Covid pandemic played a major role in the streams and songs producing as the number spiked highly in the year 2020 and 2021.
- January is the best month for releasing songs and most songs are released in January itself because of the mind openness of the people.
- At time artist peaked the charts on the single day when they passed away.
- Olivia Rodrigo came on Spotify in 2017, and her significance boomed in 2020 and 2021 because of her 2 biggest hits.
- Comedy music has recorded maximum duration in the music industry.
- Most popular genre is Children as now a days parents use the medium of music to keep their toddlers engaged.
- The causticness, loudness and energy play crucial role in audibility of the songs and the loudness of the songs have gone higher over the years.

References:

- <https://www.kaggle.com/code/liwinson/spotify-charts-artist-diversity-by-country>
- <https://www.kaggle.com/code/dhruvildave/spotify-charts-exploratory-data-analysis>
- <https://www.musicbusinessworldwide.com/bad-bunny-is-king-of-spotify-setting-new-all-time-record-for-streams-in-a-single-day/>
- <https://newsroom.spotify.com/news/>
- <https://en.wikipedia.org/wiki/Spotify>