DATA AND RESULTS VISUALIZATION PROJECT: THE IMPACT OF COVID-19 IN MUSIC INDUSTRY

Italy streams trend analysis

The first step in our analysis consists of an accurate examination of the platform streams in **Italy** using the **Spotify top 200 weekly**. The analysis is needed to gain insight into the events that cause the number of streams to increase or decrease. At first, we will plot the number of streams for the different years: *2017, 2018, 2019*, and *2020*.
We wish to identify patterns and interesting trends, such as a decrease in the number of streams due to the spread of the pandemic, or event-related peaks.

The first dataframe extracted from our dataset and taken into consideration contains four attributes:

1. `avg_streams` : is the average number of streams computed on the weekly top 200, obtained by summing the number of streams of each track in the top 200 in each week and then divided by the number of tracks (200)
2. `date` : represents the start of the week
3. `year`
4. `number_of_weeks` : Index of the week in a given year, since Spotify does not provide consistent week data.
   For different years, we could have the first week of the year starting from January 1st, while another year could start from January 3rd.
   The offset between weeks of different years is only a few days, so it does not impact our analysis.

The resulting dataframe looks as follows:

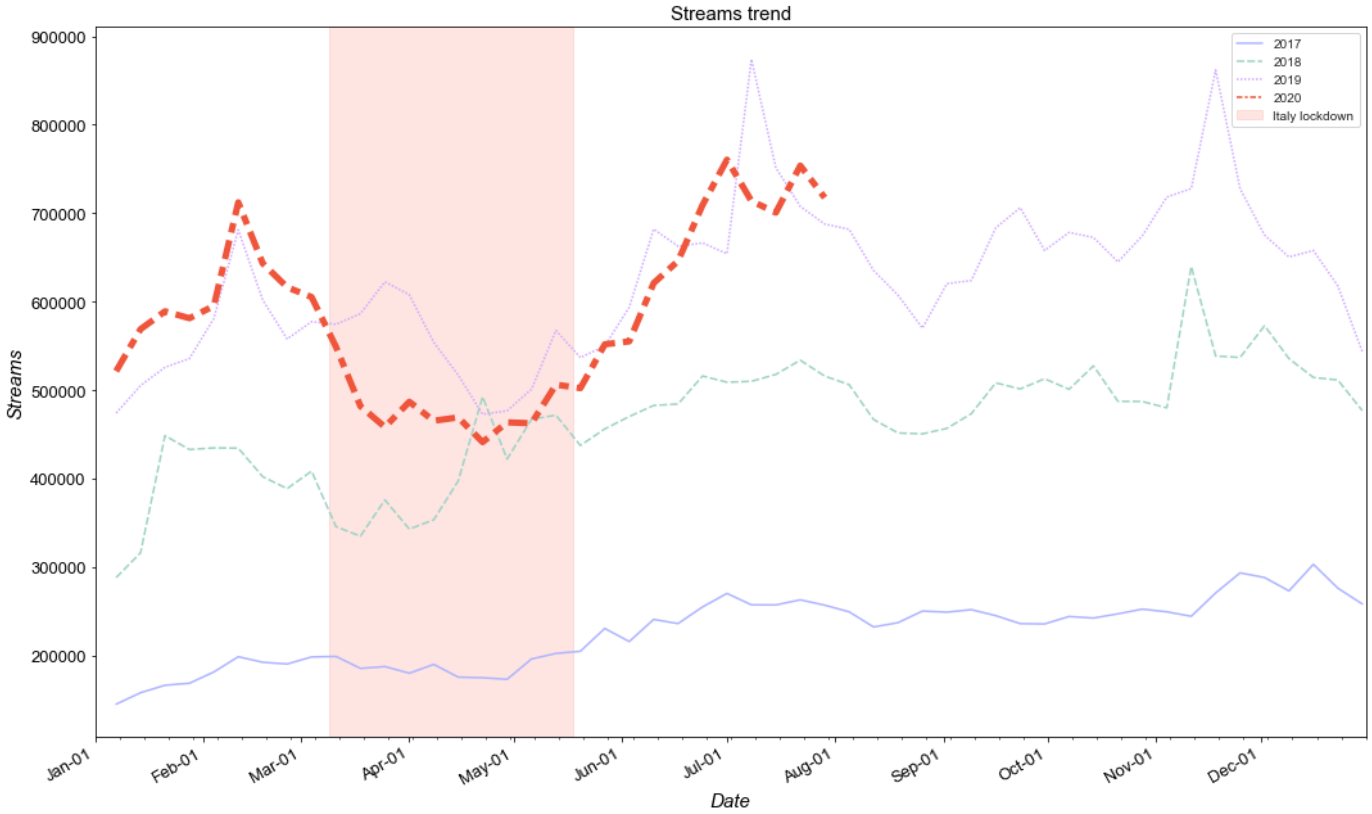
date	avg_streams	date	year	number of weeks
2017-01-06	145502.475	2017-01-06	2017	0
2017-01-13	158424.445	2017-01-13	2017	1
2017-01-20	166912.255	2017-01-20	2017	2
2017-01-27	169132.215	2017-01-27	2017	3
2017-02-03	181889.905	2017-02-03	2017	4
...	...	...	...	...
2020-06-26	760505.945	2020-06-26	2020	25
2020-07-03	713891.695	2020-07-03	2020	26
2020-07-10	700764.875	2020-07-10	2020	27
2020-07-17	753817.945	2020-07-17	2020	28
2020-07-24	717685.125	2020-07-24	2020	29

186 rows × 4 columns

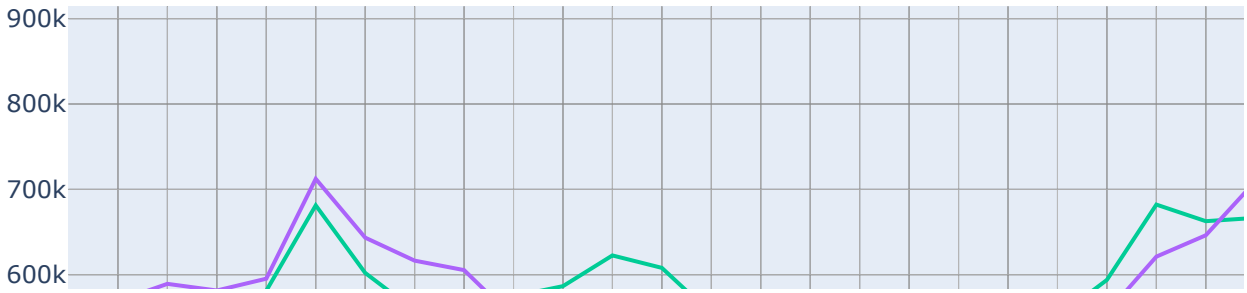
Plotting streams trend in Italy

The results below show that the beginning of the pandemic and the consequent lockdown has caused a significant decrease in the number of streams, compared to the same period in previous years.

We can also notice a constant increase in streams that seem to reach a plateau in 2020/2019. The number of streams during 2017/2018/2019 has a constant upward trend, which seems to settle down between 2020 and 2019. This trend is easily explainable, considering Spotify's growing popularity and the increasing number of total users on the platform.

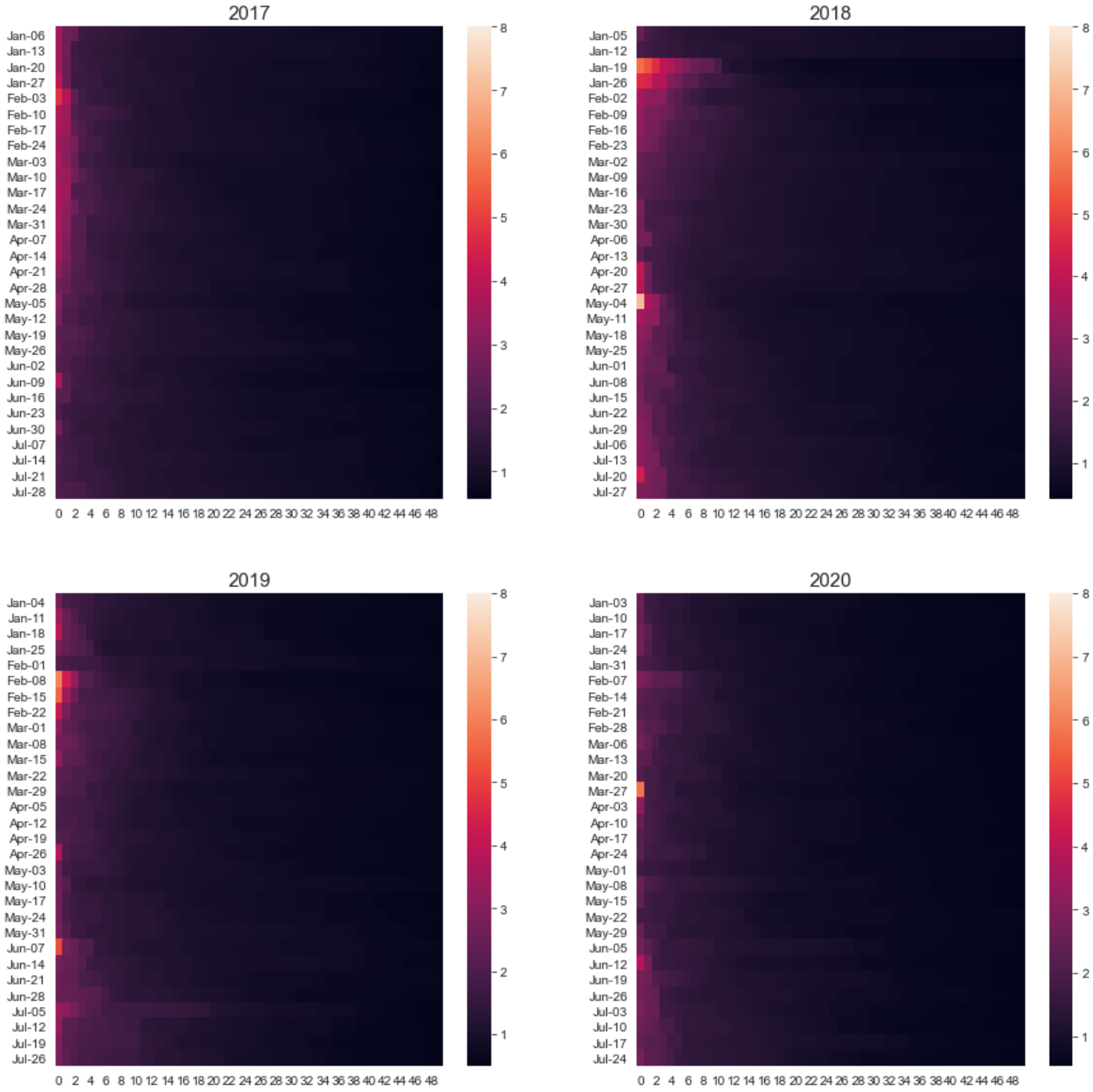


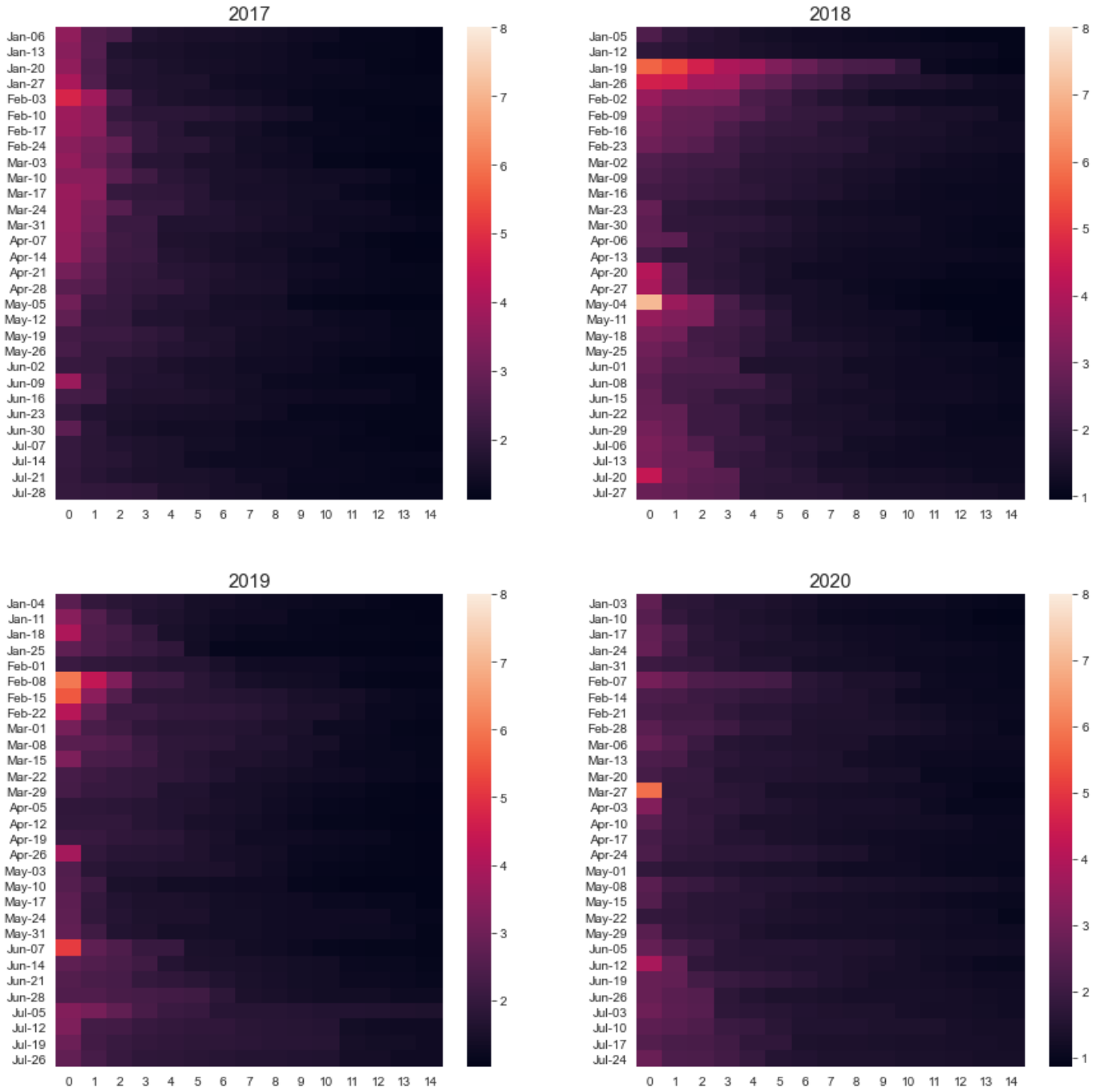
We're inserting here an interactive plot to allow the user to zoom in interesting areas of the plot.



The following heatmaps show the **distribution of the streams** per week, for each year.

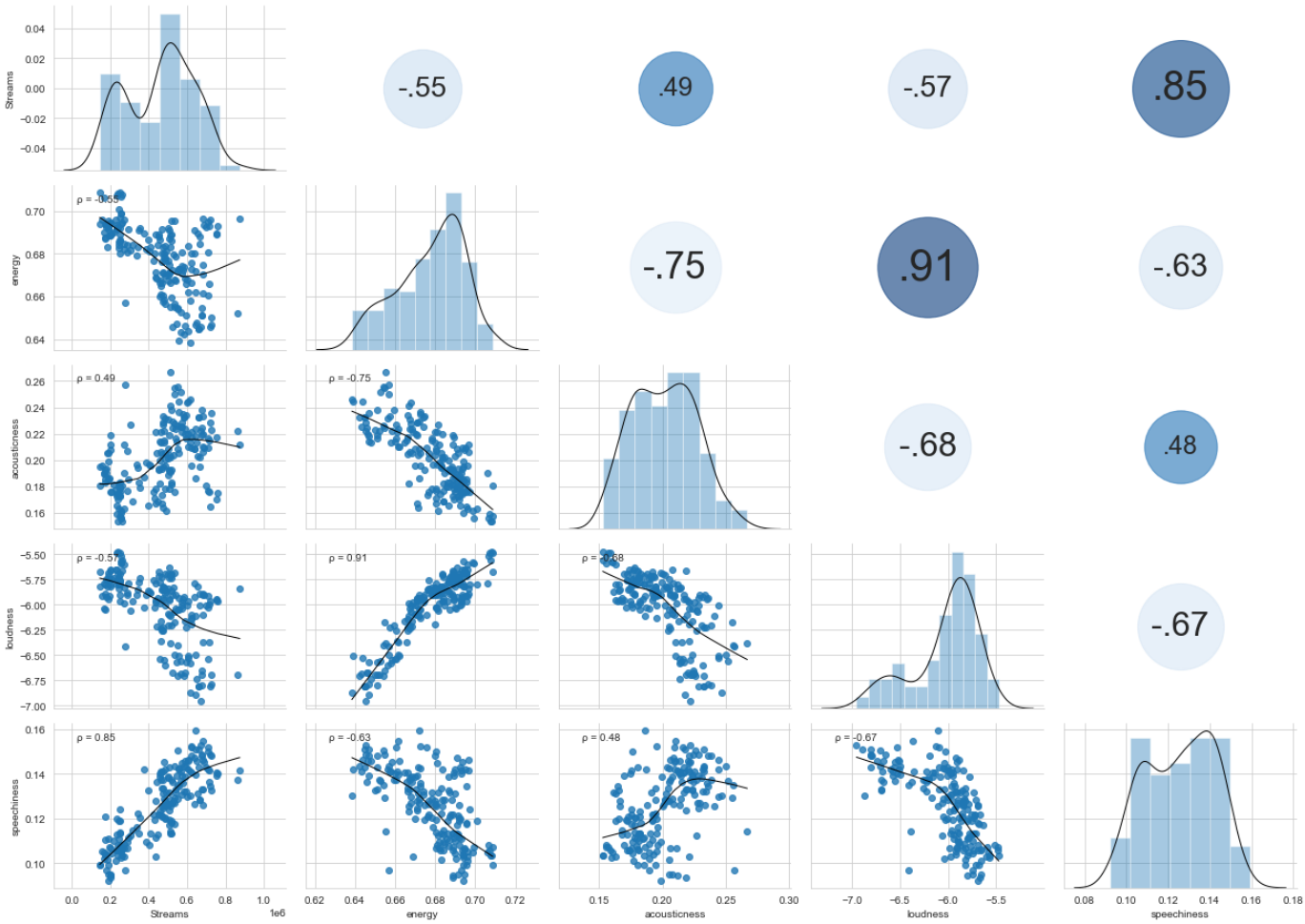
Each row is a week of the year, while each column is the top 200 position. We limited our analysis to the first 50 and 15 songs for clarity. Plotting the whole top 200 resulted in an almost completely black heatmap. From this, we can see how the distribution of the streams changes weekly, also depending on the releases of the week.
The distribution is more squeezed towards the top positions of the chart during the weeks where important releases have happened.





Musical Features Correlation

The following analysis aims to highlight emerging trends in popular music by analyzing the correlation of certain musical features, such as energy, acousticness, loudness, and speechiness, with the platform's number of streams. The high correlation between **streams** and **speechiness** is a clear indication that **Hip-Hop/Rap/Trap** tracks are dominating the Spotify's top 200.



Analyzing Stream Peaks

There are some significant peaks in the `avg_streams` that we wish to analyze to unravel the causes of such spikes. Let us retrieve the top 200 of the weeks during which those peaks occur.

7-14 February 2020 and 8-15 February 2019 peaks

We retrieve the top 200 of these weeks and analyze it.

In the following plot we're going to analyze the stream peak in the *7-14 February 2020* as highlighted on the stream trend plot (the upper one).

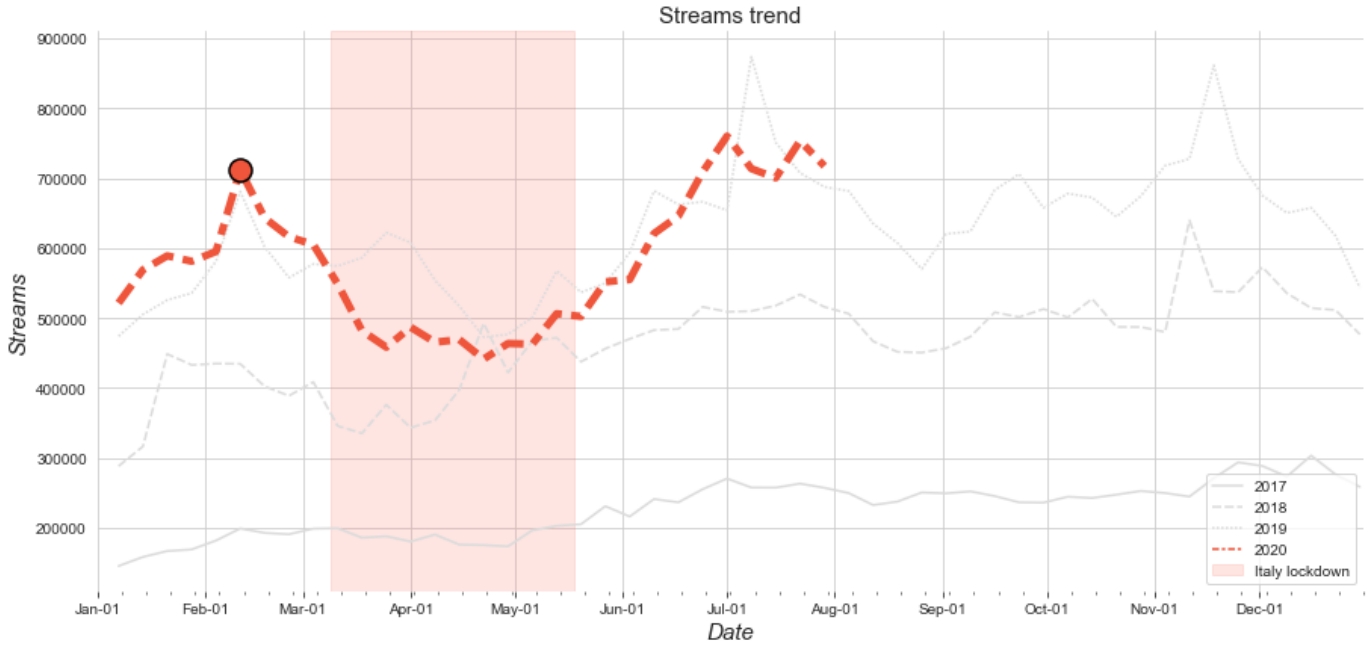
The barchart shows the number of songs in the top 200 per artist, while the treemap shows the percentage of streams per artist.

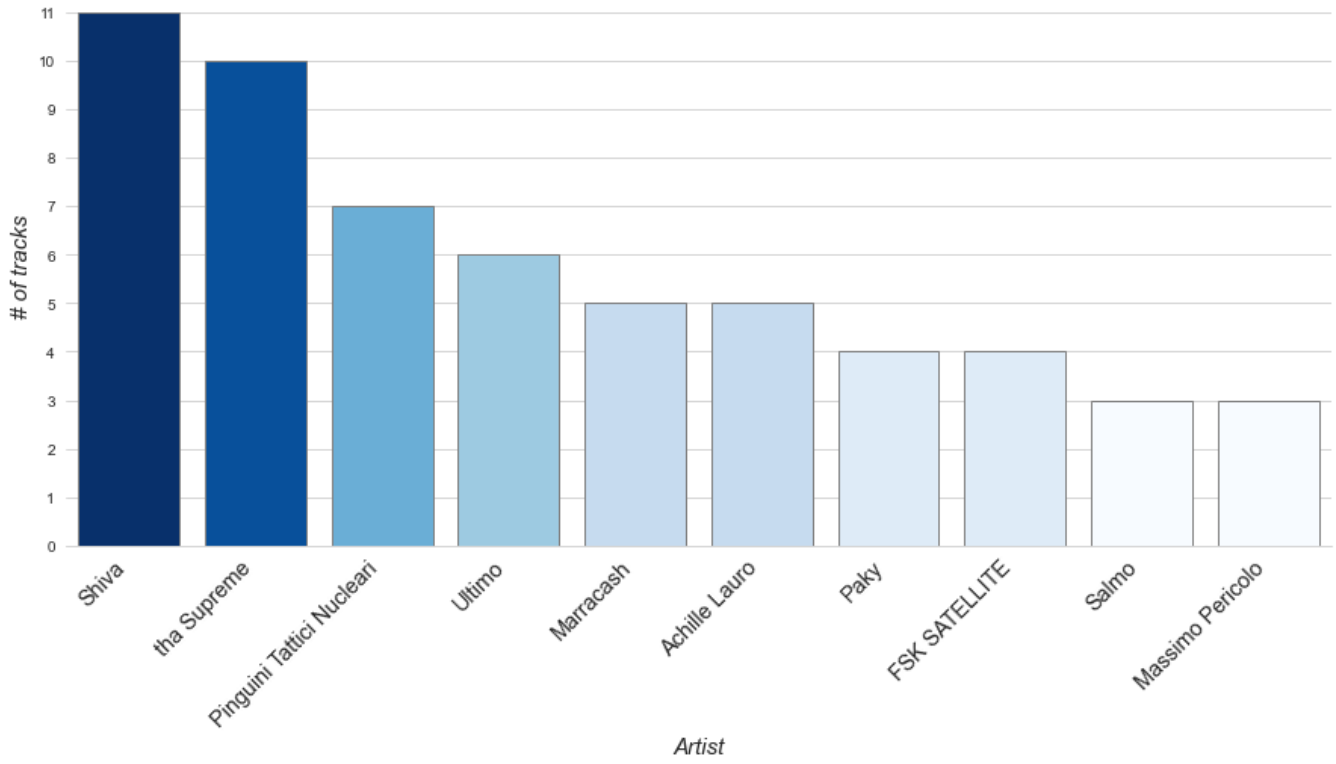
The barchart shows a major number of tracks by ***Shiva***. He released the ***Routine*** EP on the **31 January of 2020**.
However, the barchart can be misleading.
There has indeed been a major release by ***Shiva*** (we will see later how releases impact streams), but in this case it's not the major reason for streams increment.

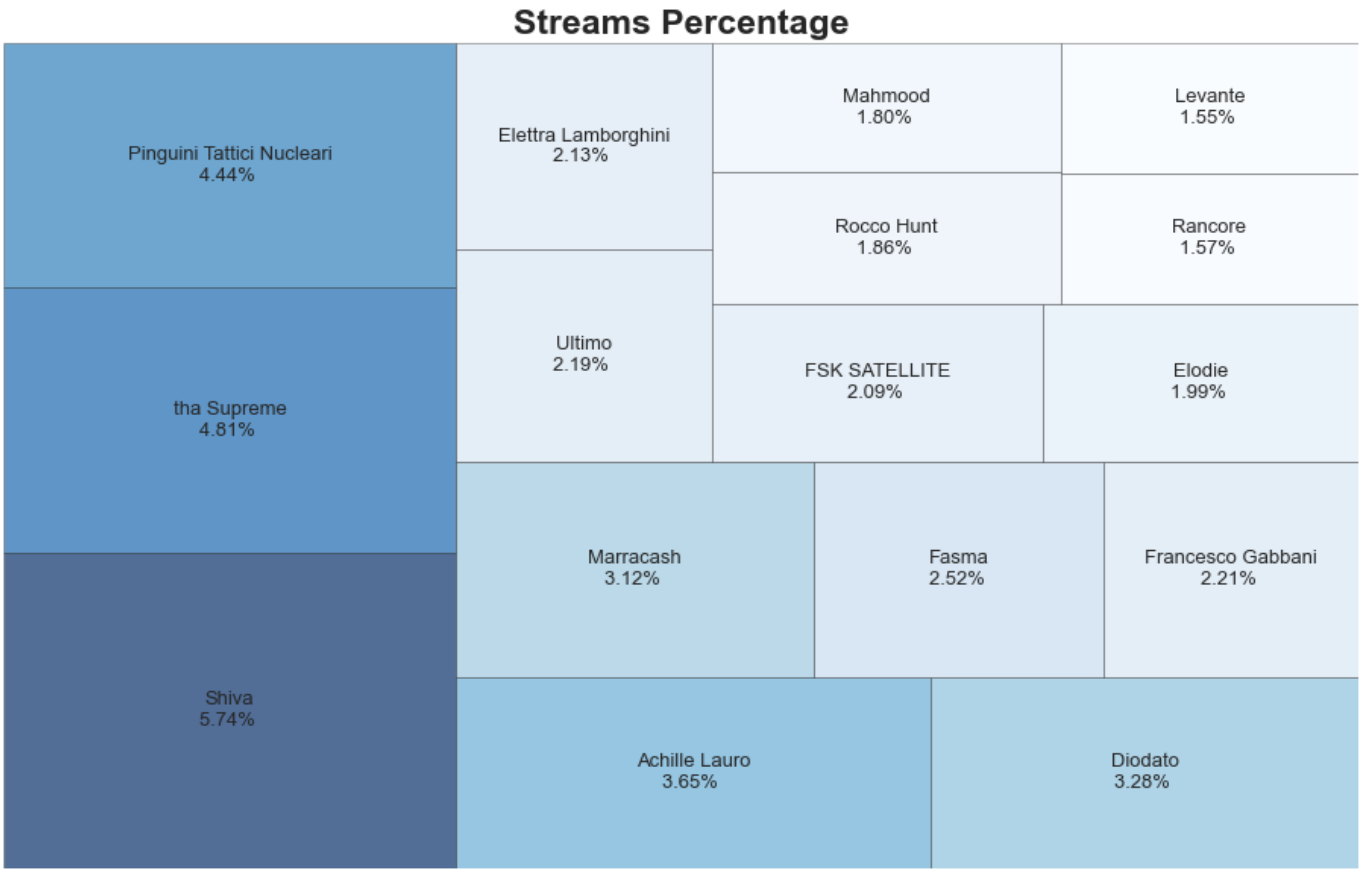
In fact, if we analyze the treemap, we can notice that there are many artists who have a good percentage of streams even if their number of tracks in the top 200 is less or equal to three.

Such artists are ***Diodato***, ***Marracash***, ***Fasma***, ***Francesco Gabbani*** and ***Elettra Lamborghini***.

The plots show also a major presence of ***ThaSupreme***, his presence is given by it's major release in *November 2019*, which guaranteed him a long presence on the top charts (we will analyze his case later)







The following result should make ourselves ask: *Why are there some artist like Diodato who have a quite high percentage of streams even if they have three or less tracks in the top 200?*

As we might notice from the following top 10, the first top 8 songs come from **Sanremo Festival**, with ***Diodato*** in first position, which happens to be the winner of **Sanremo 2020**.

	Track Name	Artist	Streams	genre
0	Fai rumore	Diodato	4268423	Pop
1	Ringo Starr	Pinguini Tattici Nucleari	3840527	Pop
2	Me ne frego	Achille Lauro	3309741	Hip-Hop/Rap
3	Per sentirmi vivo	Fasma	3192031	Hip-Hop/Rap
4	Viceversa	Francesco Gabbani	3144858	Pop
5	Musica (E Il Resto Scompare)	Elettra Lamborghini	3035842	Urbano latino
6	Andromeda	Elodie	2380992	Hip-Hop/Rap
7	Tikibombom	Levante	2215022	Pop
8	blun7 a swishland	tha Supreme	2010716	Hip-Hop/Rap
9	Boogieman (feat. Salmo)	Ghali	2001335	Hip-Hop/Rap
10	Eden (feat. Dardust)	Rancore	1892951	Hip-Hop/Rap
11	Rapide	Mahmood	1606138	Pop
12	Calmo - feat. tha Supreme	Shiva	1554386	none
13	ANSIA NO	FSK SATELLITE	1553983	Hip-Hop/Rap
14	Rosso di rabbia	Anastasio	1446845	Hip-Hop/Rap
15	Dov'è	Le Vibrazioni	1418915	Pop

	Track Name	Artist	Streams	genre
16	Dance Monkey	Tones And I	1372288	Alternative
17	Vai Bene Così	Leo Gassmann	1305747	Pop
18	fuck 3x	tha Supreme	1292623	Hip-Hop/Rap
19	Ti volevo dedicare (feat. J-AX & Boomdabash)	Rocco Hunt	1276149	Hip-Hop/Rap

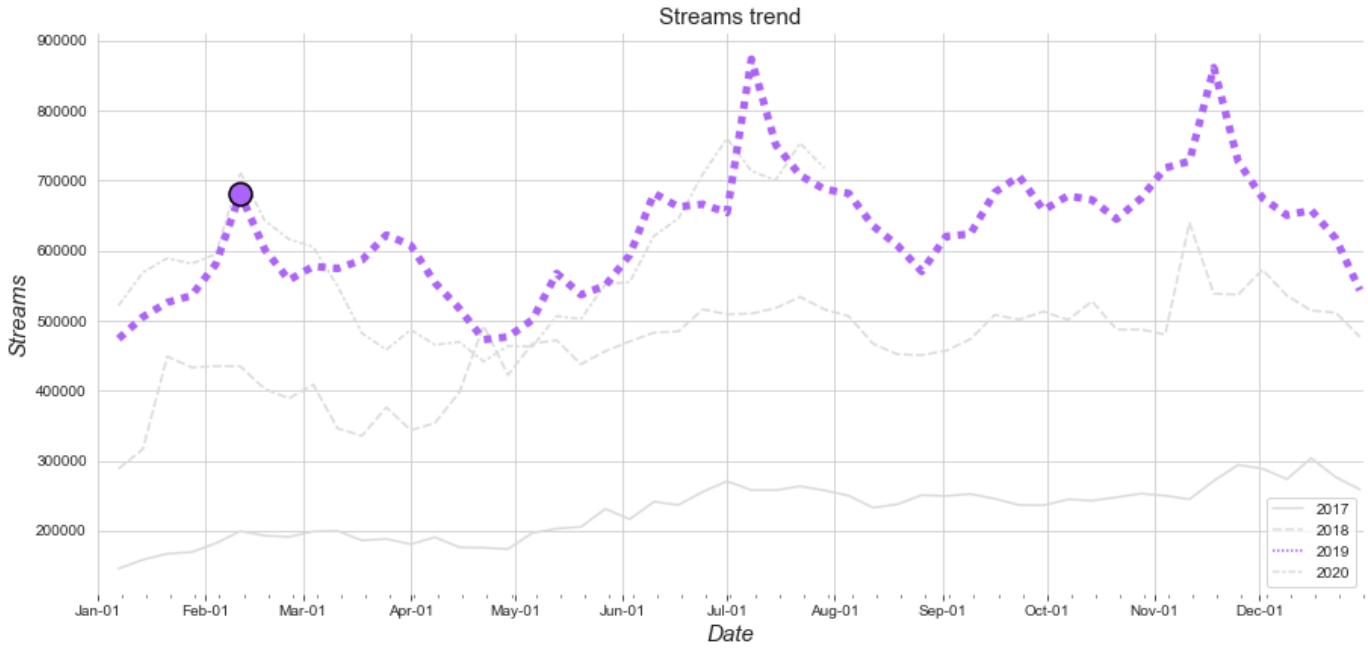
Now we're going to analyze the same week, but in 2019. We're analyzing the period *8-15 February 2019*

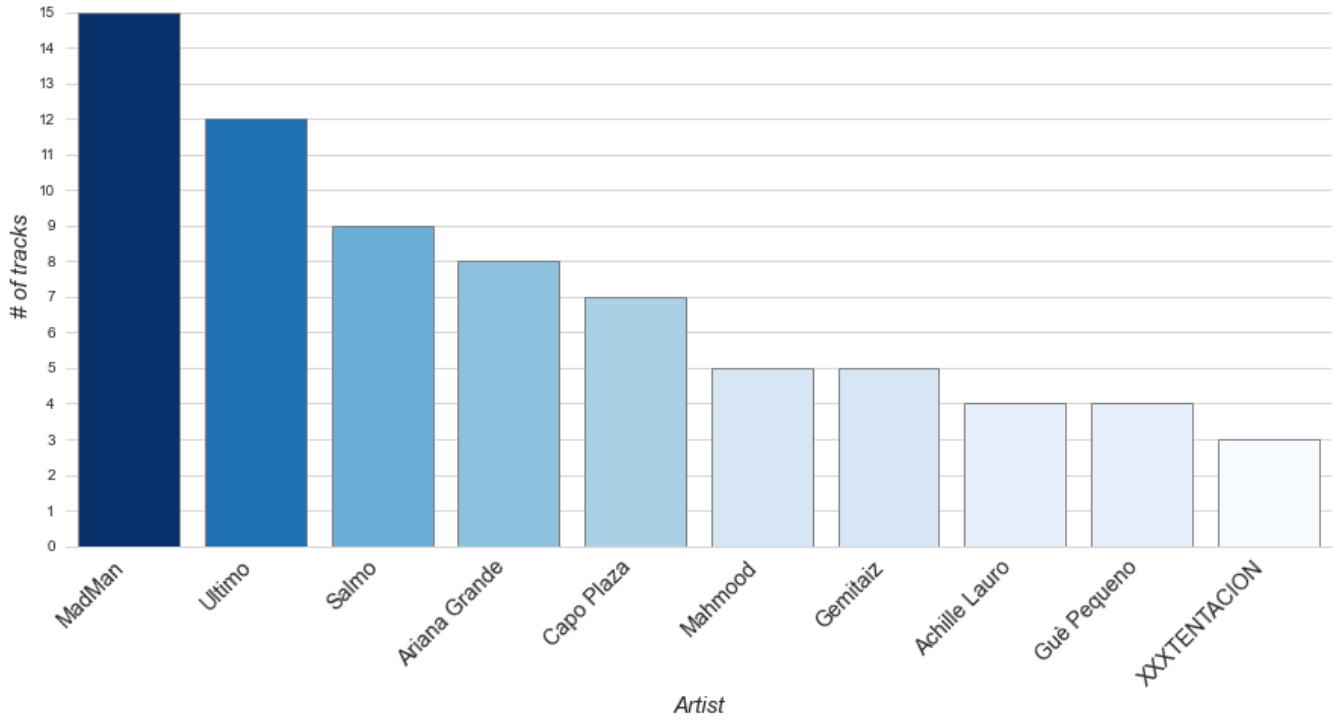
The histogram on the left shows a major number of tracks by ***Madman***. He released the ***MM Vol.3 Mixtape*** on the **1 February of 2019**.
However, if we look at the histogram containing the percentage of streams, we can see that ***Madman*** is only in third position, with ***Ultimo*** and ***Mahmood*** in first and second position respectively.

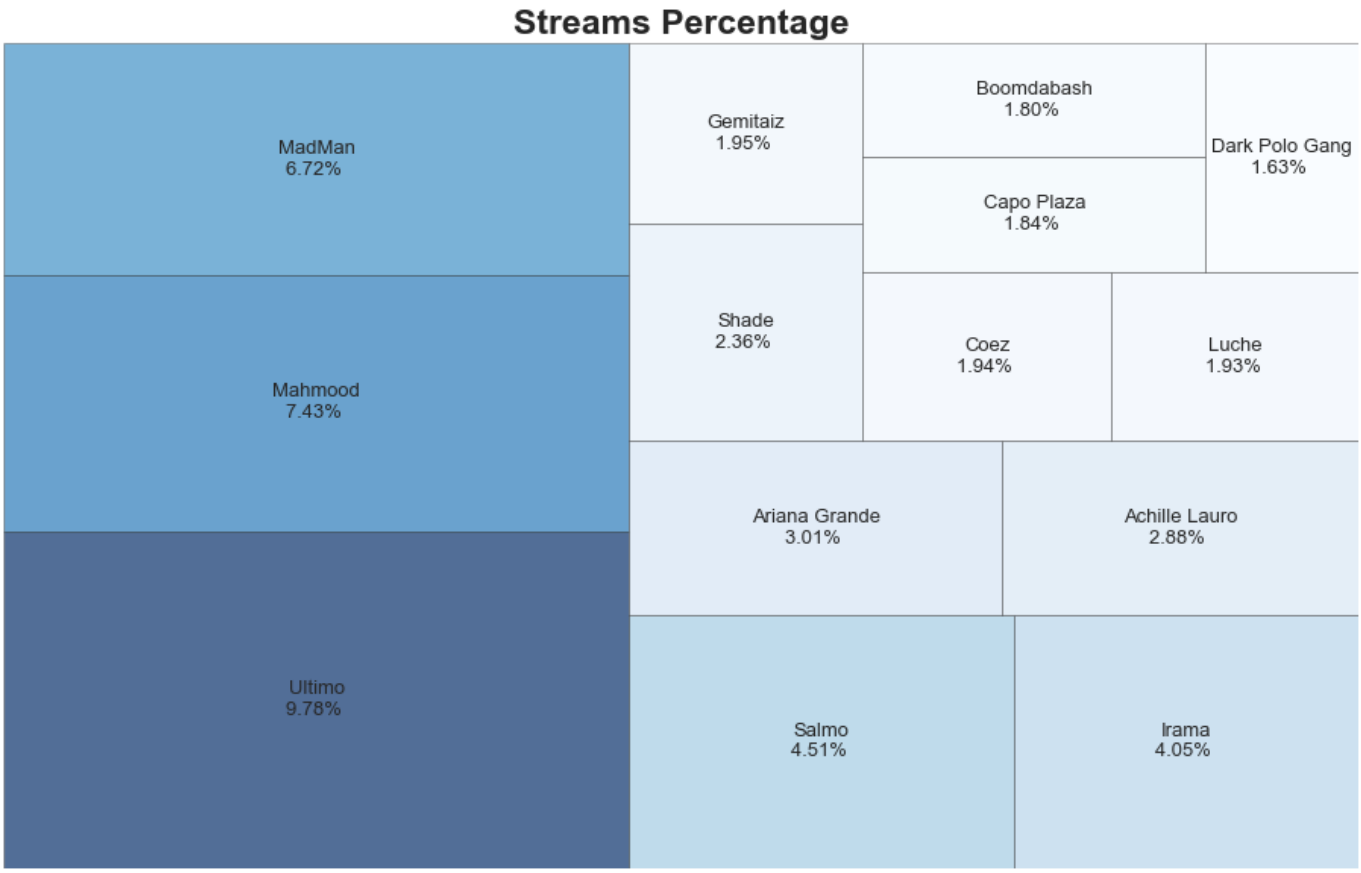
Mahmood won the **Sanremo Festival 2019** while ***Ultimo*** ended up taking the second place.

There has also been a lot of discussion and criticisms about ***Mahmood*** first place. In fact, the favourite winner by the public was ***Ultimo*** with the **48,80%** of votes, while ***Mahmood*** got only the **20,95%** (source: https://www.sorrisi.com/musica/sanremo/sanremo-2019-analisi-del-televoto-e-dati-delle-giuri/).

The technical jury decided to assign the first place to ***Mahmood***, this generated a lot of criticisms and discussion on public media, which guaranteed ***Ultimo*** a huge visibility and streams on Spotify.


Streams trend
Streams
900000
800000
700000
600000
500000
400000
300000
200000
Jan-01
Feb-01
Mar-01
Apr-01
May-01
Jun-01
Jul-01
Aug-01
Sep-01
Oct-01
Nov-01
Dec-01
Date
2017
2018
2019
2020


# of tracks
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
MadMan
Ultimo
Salmo
Ariana Grande
Capo Plaza
Mahmood
Gemitaiz
Achille Lauro
Guè Pequeno
XXXTENTACION
Artist



If we plot the top 200 weekly dataset, we can notice that, out of the top 10 songs, 6 of them are from *Sanremo Festival*.

	Track Name	Artist	Streams	genre
0	Soldi	Mahmood	8223614	Pop
1	I Tuoi Particolari	Ultimo	5902025	Pop
2	La ragazza con il cuore di latta	Irama	4458947	Pop
3	Senza farlo apposta	Shade	2919684	Hip-Hop/Rap
4	Rolls Royce (feat. Boss Doms & Frenetik&Orang3)	Achille Lauro	2848988	Hip-Hop/Rap
5	Per Un Milione	Boomdabash	2456164	Pop
6	È sempre bello	Coez	2216325	Pop
7	IL CIELO NELLA STANZA (feat. NSTASIA)	Salmo	2001997	Hip-Hop/Rap
8	Sweet but Psycho	Ava Max	1884427	Pop
9	Calma - Remix	Pedro Capó	1859418	Pop Latino

Correlation between musical events and Spotify streams

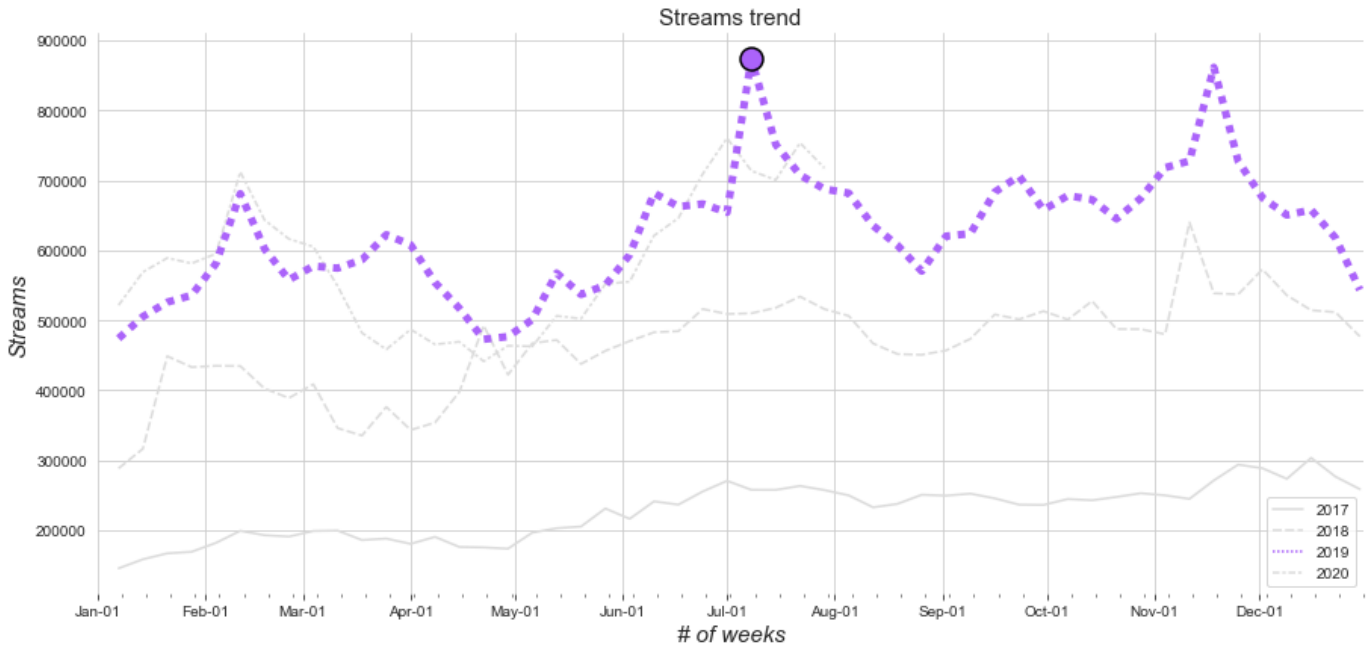
It is indeed clearer how the streams spiking up in *February* in *Italy* are generally due to **new releases** and, more importantly, to **Sanremo Festival**. This analysis reveals the correlation between musical events (festivals, awards, etc.) and stream trends that will be further confirmed in subsequent analysis.

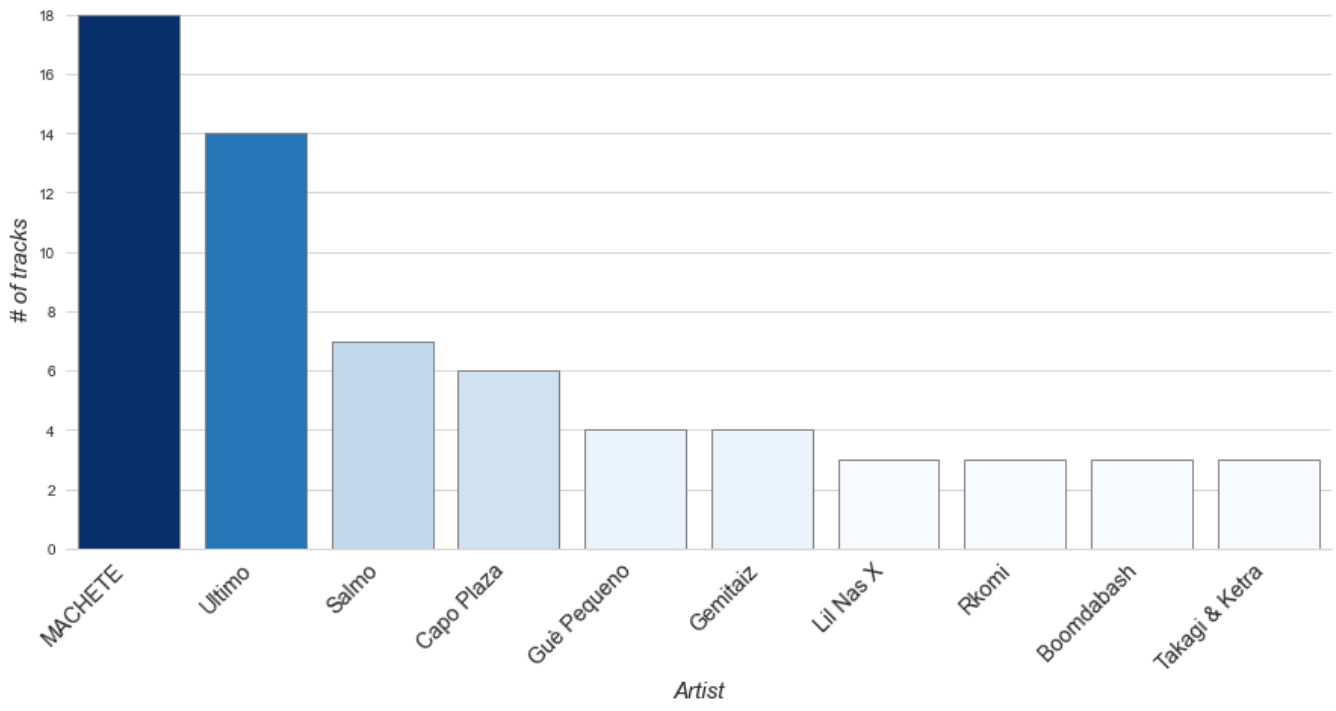
5 - 12 July 2019

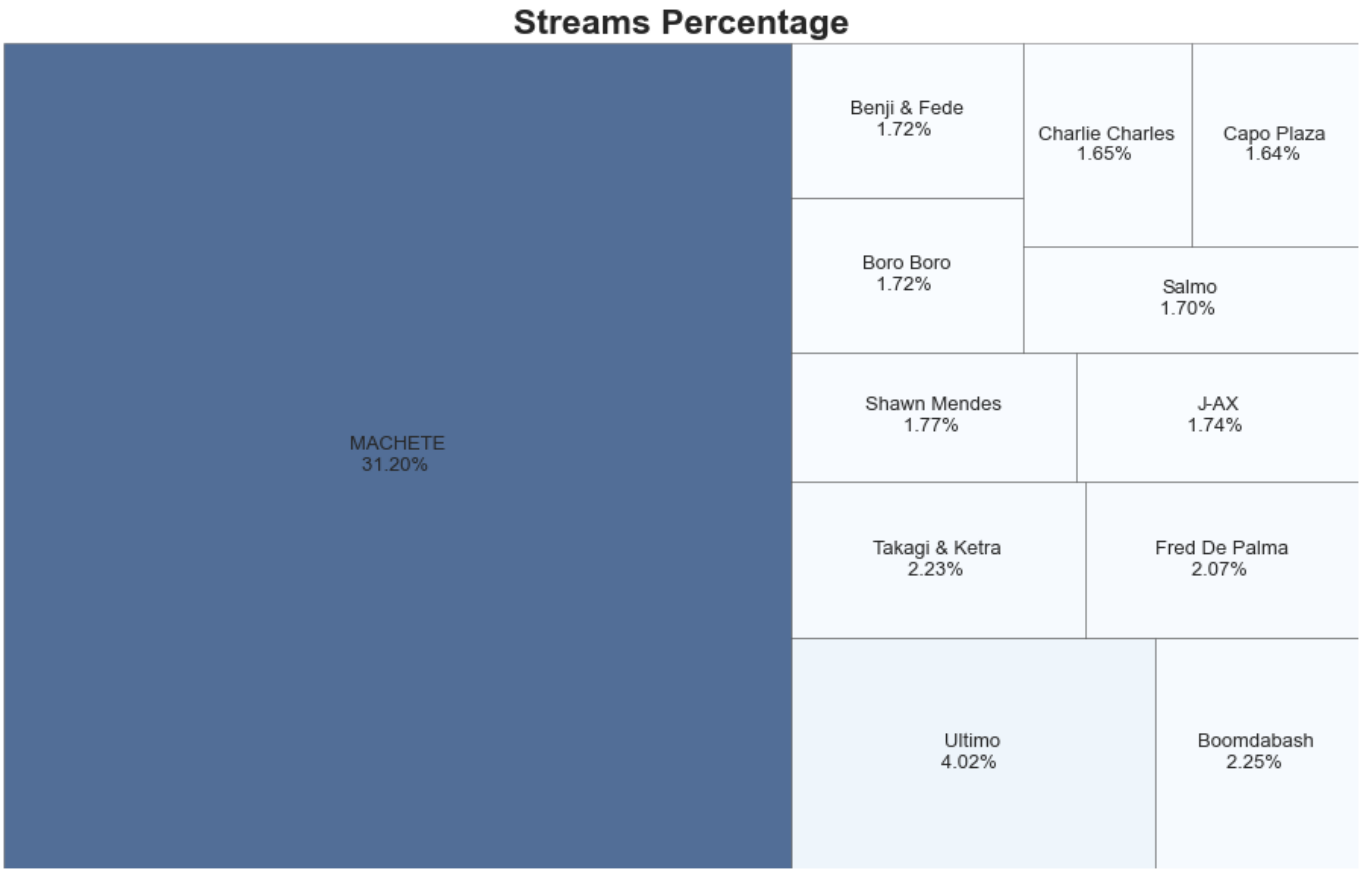
We can notice a huge spike in the number of streams on this week.

We can notice there's a big number of track from ***MACHETE***. In fact on the 5 july 2019 the ***MACHETE Mixtape 4*** has been released, which had a huge success, with major collaborations between big rappers in the italian music scene.

Also, more than the **30%** of the weekly streams are given by ***MACHETE***.





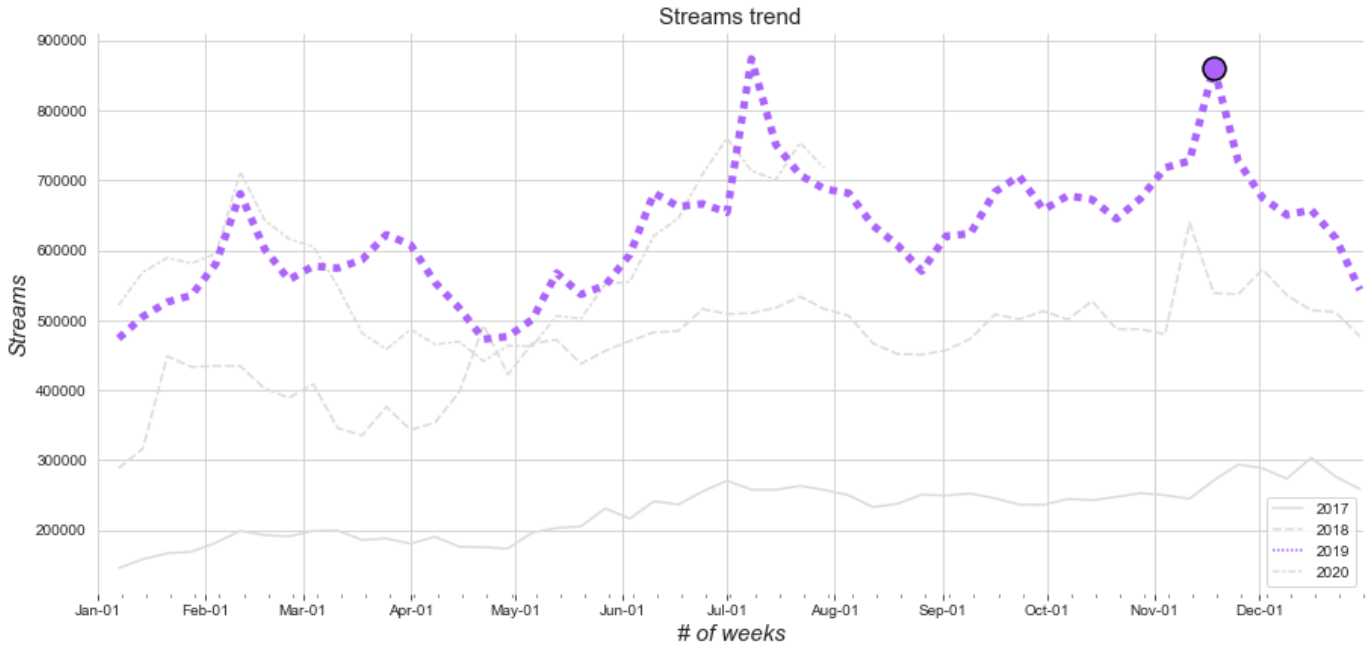


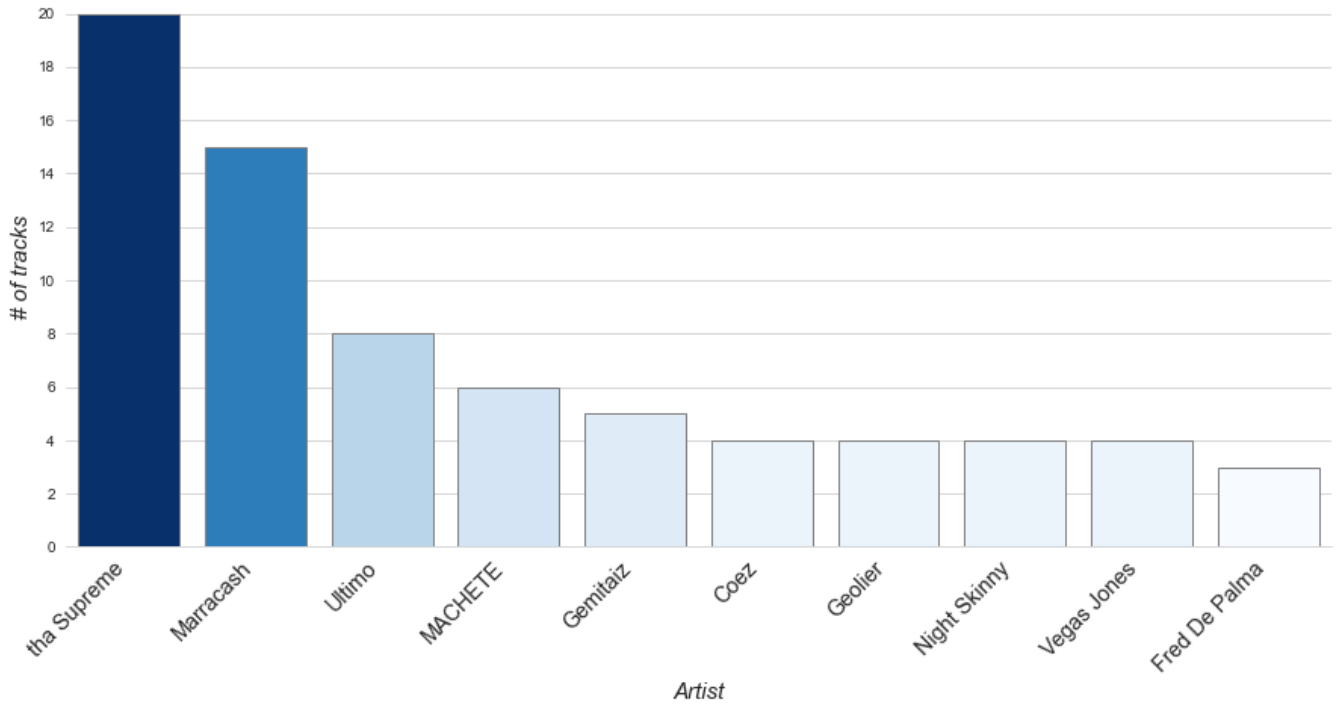
We can also notice that almost the whole top 10 comes from that album.

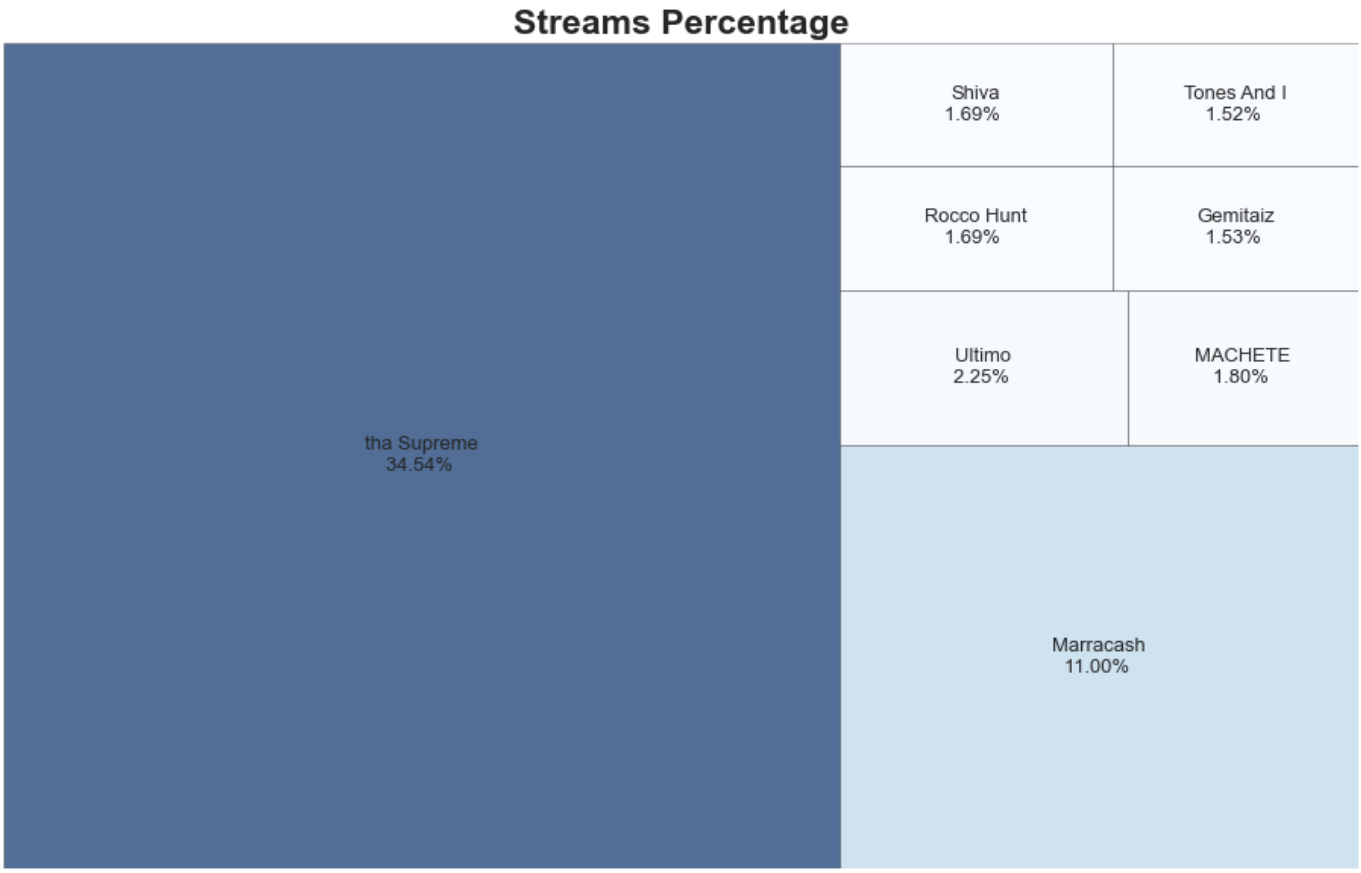
	Track Name	Artist	Streams
0	YOSHI (feat. Fabri Fibra) - prod. Strage	MACHETE	5816357
1	HO PAURA DI USCIRE 2 - prod. Mace	MACHETE	5489578
2	MARYLEAN (feat. Marracash) - prod. Low Kidd	MACHETE	4829415
3	STAR WARS - prod. Young Miles - Crookers x Nic...	MACHETE	4098318
4	MAMMASTOMALE (feat. Salmo) - prod. Dade	MACHETE	3680698
5	NO WAY (feat. Nitro) - prod. Low Kidd	MACHETE	3555978
6	IO PUO' (feat. Salmo) - prod. Low Kidd	MACHETE	3094979
7	Una volta ancora (feat. Ana Mena)	Fred De Palma	3056237
8	JAMBO	Takagi & Ketra	3054732
9	Dove e quando	Benji & Fede	3008972

15-22 November 2019

In the left histogram we can notice 20 songs from ***Tha Supreme***. He released highly anticipated debut album ***236451*** on the *15 November 2019*, which had a huge success.
Almost the **35%** of the weekly top 200 streams has been generated by ***Tha Supreme*** followed by ***Marracash*** who release his album ***Persona*** the previous week.


Streams trend
900000
800000
700000
600000
500000
400000
300000
200000
Streams
Jan-01
Feb-01
Mar-01
Apr-01
May-01
Jun-01
Jul-01
Aug-01
Sep-01
Oct-01
Nov-01
Dec-01
# of weeks
2017
2018
2019
2020


20
18
16
14
12
10
8
6
4
2
0
# of tracks
tha Supreme
Marracash
Ultimo
MACHETE
Gemitaiz
Coez
Geolier
Night Skinny
Vegas Jones
Fred De Palma
Artist

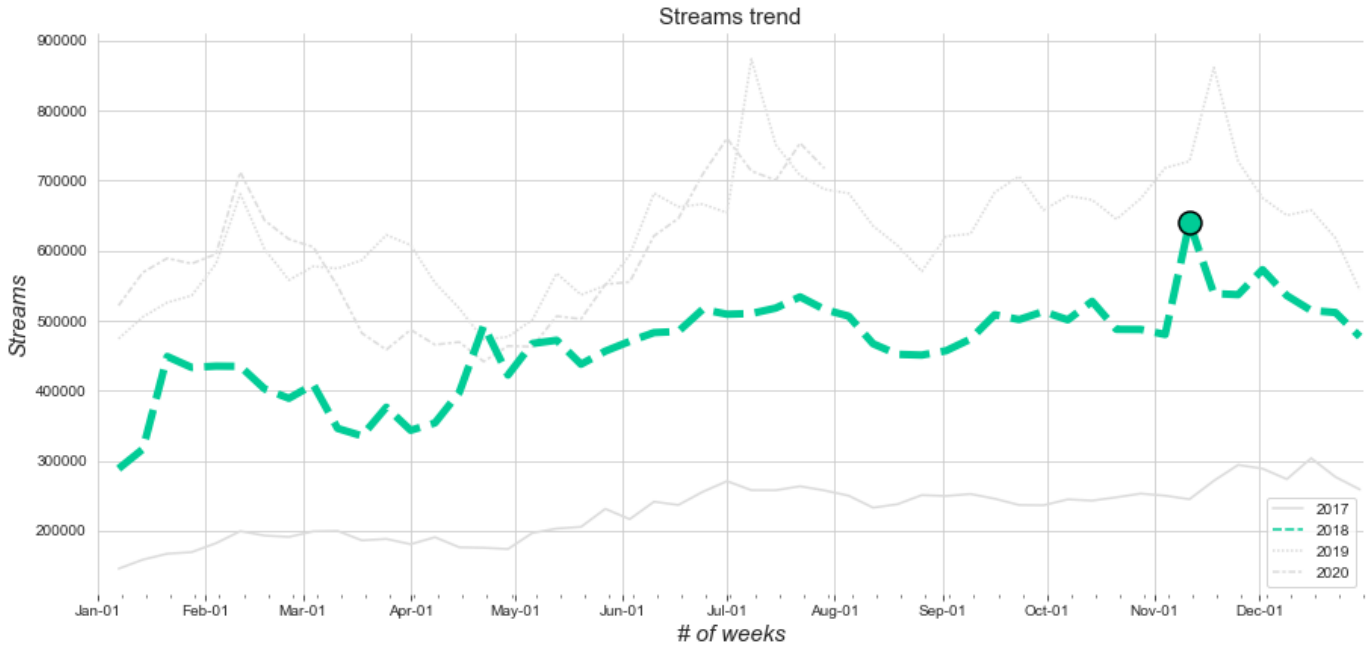


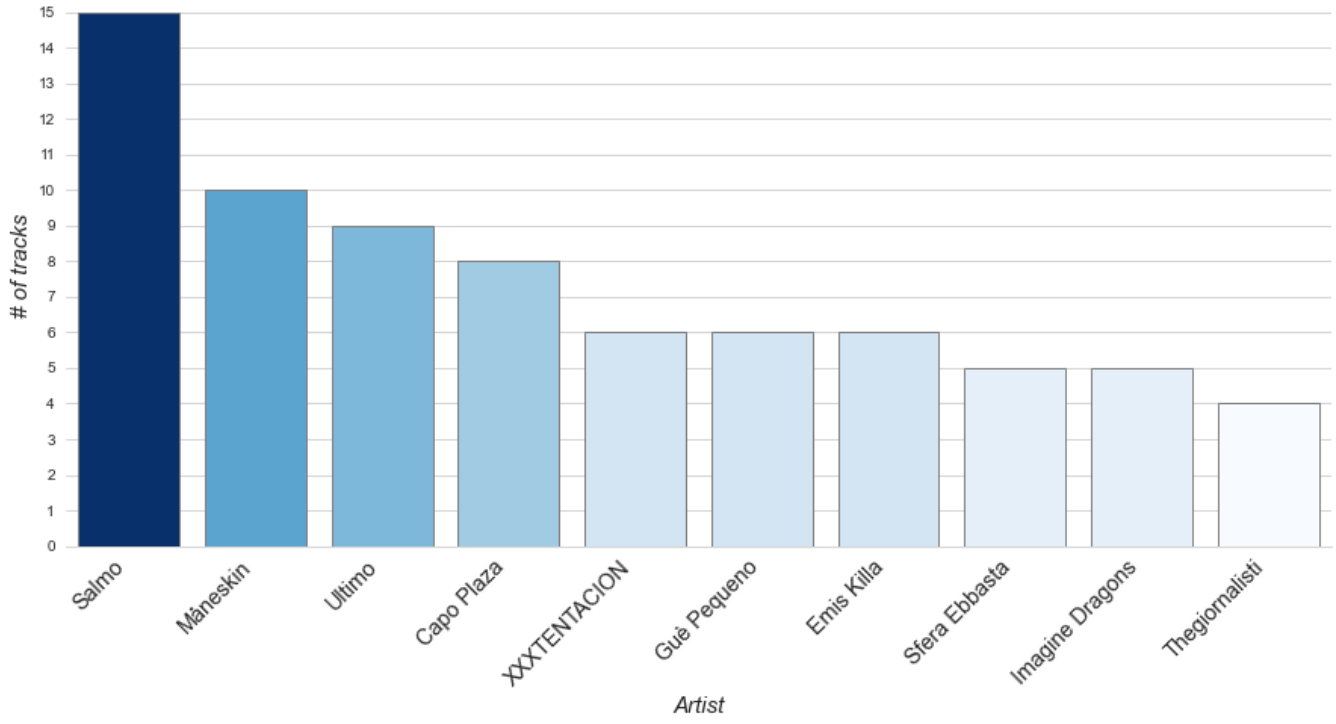
As we can see, basically the whole top 10 is by ***Tha Supreme***

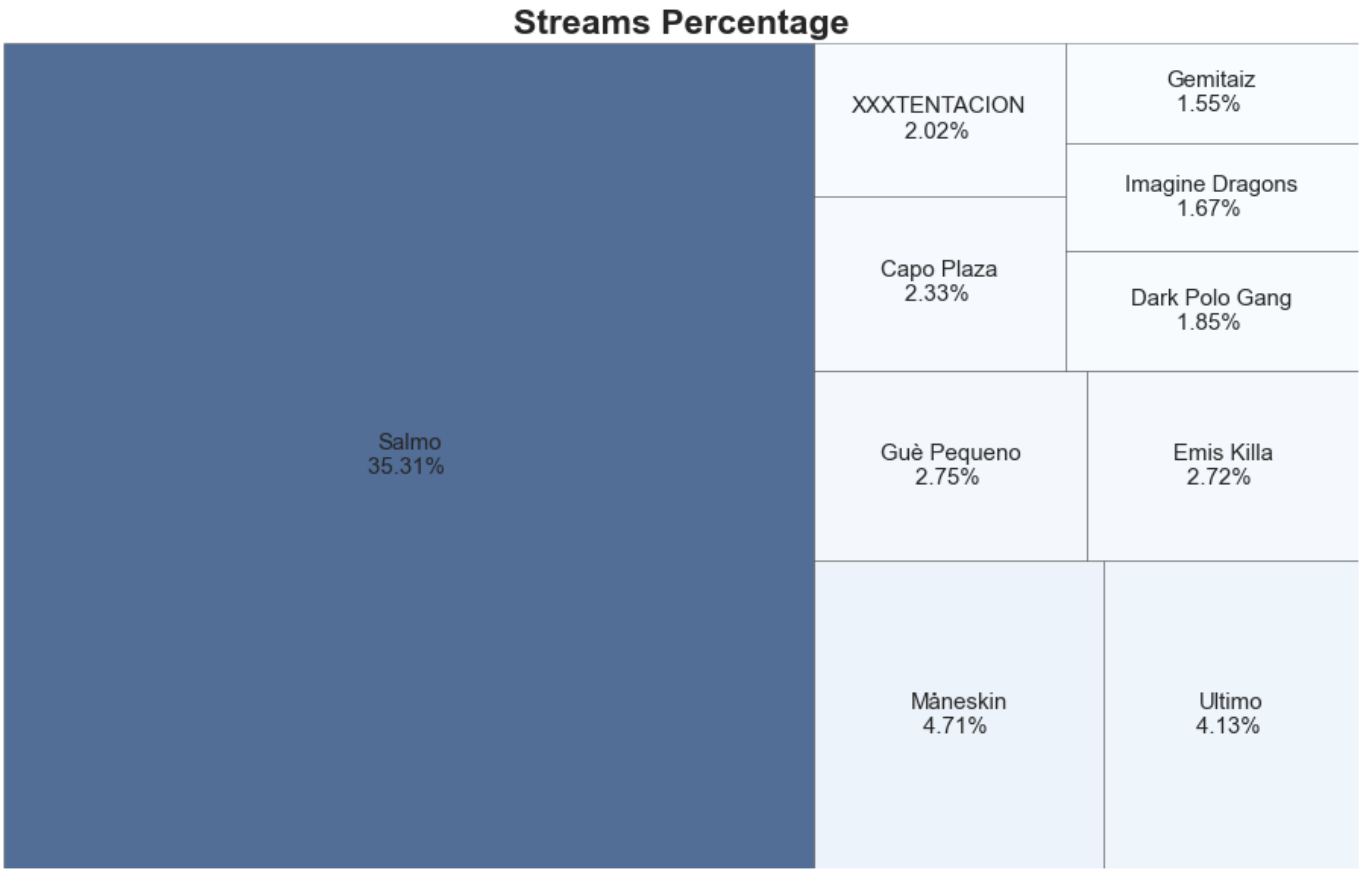
	Track Name	Artist	Streams
0	blun7 a swishland	tha Supreme	7429539
1	fuck 3x	tha Supreme	4668644
2	sw1n6o - feat. Salmo	tha Supreme	3954844
3	m12ano - feat. Mara Sattei	tha Supreme	3555240
4	SUPREME - L'ego (feat. tha Supreme & Sfera Ebb...	Marracash	3381296
5	occh1 purpl3 - feat. Marracash	tha Supreme	3329403
6	no14 - feat. Dani Faiv	tha Supreme	3081602
7	come fa1	tha Supreme	3019409
8	ch1 5ei te	tha Supreme	3012034
9	2ollipop	tha Supreme	2901181

9-16 november 2018

15 tracks by ***Salmo*** are present in the top 200. In fact his album ***Playlist*** has been released on the *9 November 2018*


Streams trend
900000
800000
700000
600000
500000
400000
300000
200000
Streams
Jan-01
Feb-01
Mar-01
Apr-01
May-01
Jun-01
Jul-01
Aug-01
Sep-01
Oct-01
Nov-01
Dec-01
# of weeks
2017
2018
2019
2020


15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
# of tracks
Salmo
Måneskin
Ultimo
Capo Plaza
XXXTENTACION
Guè Pequeno
Emis Killa
Sfera Ebbasta
Imagine Dragons
Thegiornalisti
Artist

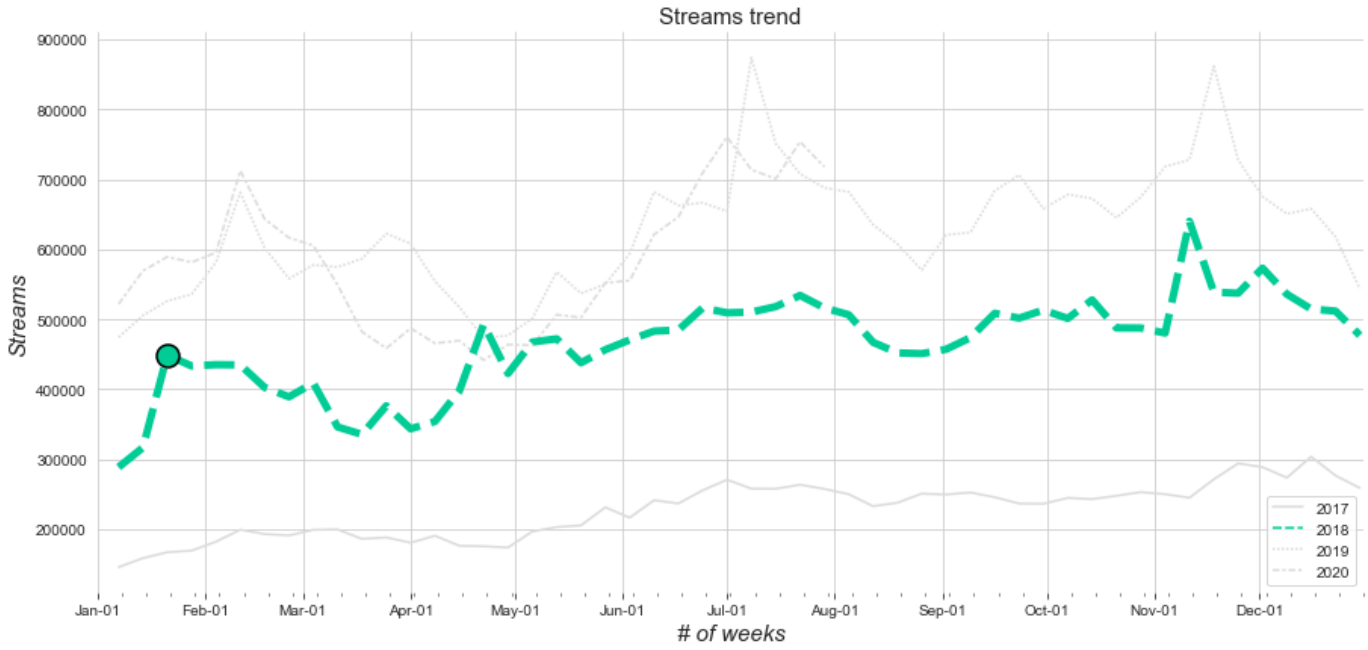


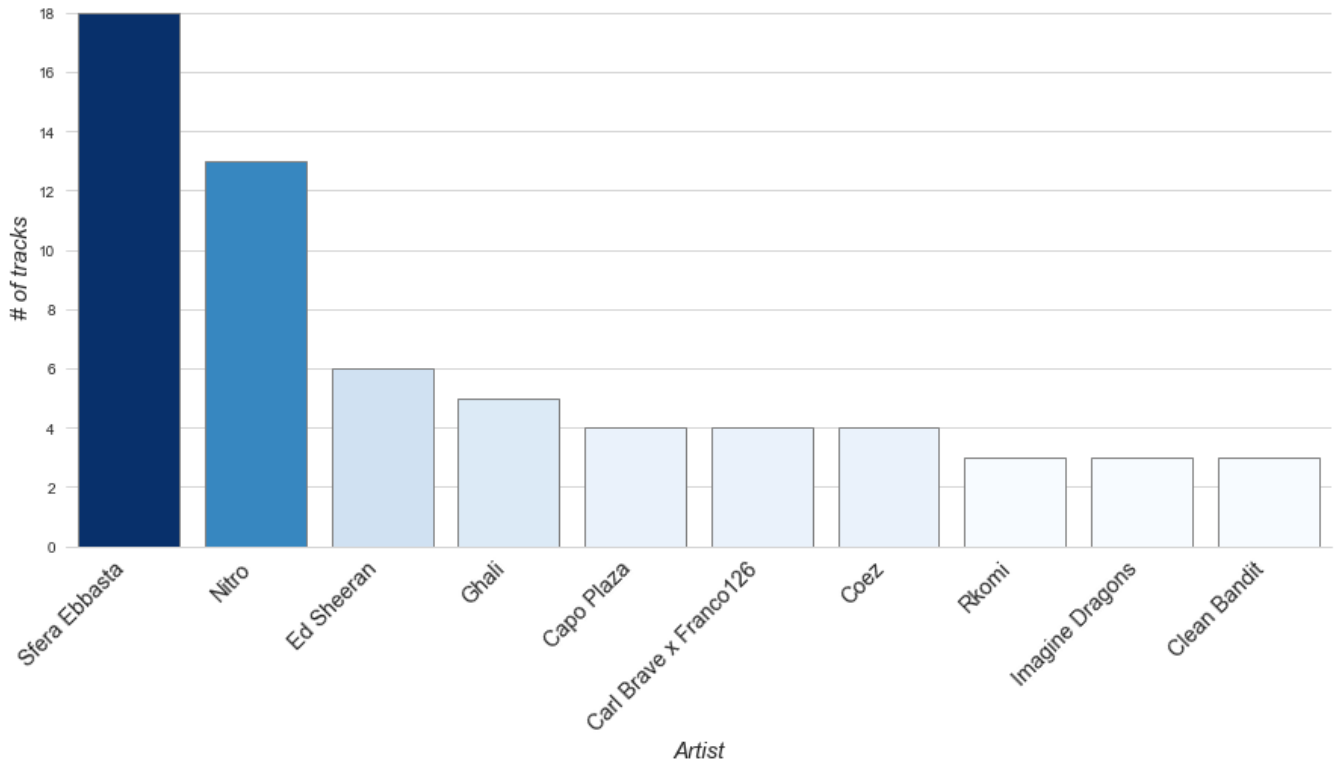
We can easily see that almost the whole top 10 is taken by ***Salmo***.

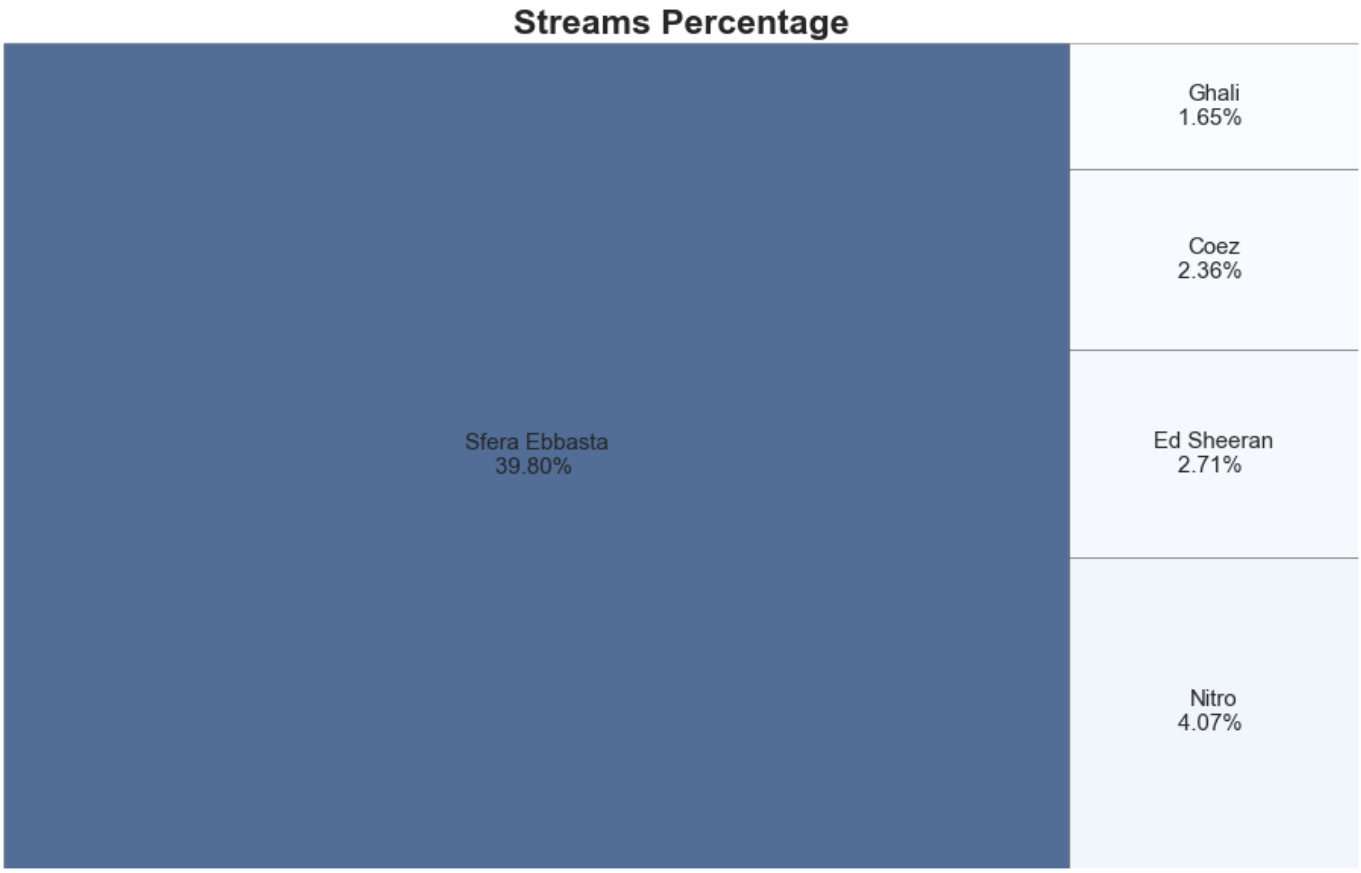
	Track Name	Artist	Streams
0	CABRIOLET (feat. Sfera Ebbasta)	Salmo	5434298
1	STAI ZITTO (feat. Fabri Fibra)	Salmo	5195708
2	IL CIELO NELLA STANZA (feat. NSTASIA)	Salmo	5163645
3	LUNEDI'	Salmo	3859383
4	HO PAURA DI USCIRE	Salmo	3493788
5	SPARARE ALLA LUNA (feat. Coez)	Salmo	3475966
6	RICCHI E MORTI	Salmo	3390260
7	90MIN	Salmo	3216286
8	DISPOVERY CHANNEL (feat. Nitro)	Salmo	3118643
9	Torna a casa	Måneskin	2797833

19-26 January 2018

We can easily spot 18 tracks by ***Sfera Ebbasta***. He released his album ***Rockstar*** on the *19 Janruary 2018*, which has been a huge success in the italian trap scene.


Streams trend
900000
800000
700000
600000
500000
400000
300000
200000
Streams
Jan-01
Feb-01
Mar-01
Apr-01
May-01
Jun-01
Jul-01
Aug-01
Sep-01
Oct-01
Nov-01
Dec-01
# of weeks
2017
2018
2019
2020


18
16
14
12
10
8
6
4
2
0
# of tracks
Sfera Ebbasta
Nitro
Ed Sheeran
Ghali
Capo Plaza
Carl Brave x Franco126
Coez
Rkomi
Imagine Dragons
Clean Bandit
Artist



The whole top 10 is taken by ***Sfera Ebbasta***.

	Track Name	Artist	Streams
0	Cupido (feat. Quavo)	Sfera Ebbasta	5105410
1	Rockstar	Sfera Ebbasta	4722530
2	Sciroppo (feat. DrefGold)	Sfera Ebbasta	4156428
3	Ricchi x Sempre	Sfera Ebbasta	3552712
4	Serpenti A Sonagli	Sfera Ebbasta	3381143
5	Uber	Sfera Ebbasta	2894123
6	XNX	Sfera Ebbasta	2552249
7	Leggenda	Sfera Ebbasta	2216057
8	20 Collane	Sfera Ebbasta	2071546
9	Bancomat	Sfera Ebbasta	2051597

Releases and events impact number of streams

To conclude, we can observe a correlation between the number of streams on Spotify and major releases and musical events.
This correlation becomes more evident than before in the last two years, with Spotify becoming a widespread platform used daily by its users. In 2017 and the beginning of 2018, the service's popularity was still growing, as we can see by the constant increase in the number of streams.

Limited to Italy, most spikes were given by Italian **Hip-Hop/Rap/Trap** artists, as we have already anticipated with our correlation plot, which showed a high correlation between speechiness and

streams.

These charts reported the analysis only of a limited number of peaks for sake of clarity. The analysis has been performed on all peaks in the plot.

We also identified a decreasing trend in the number of streams during quarantine/pandemic. Considering the previous results, one should naturally ask himself: "*Is this decreasing trend caused directly by the pandemic, or is it caused by the lack of new releases and musical events?*"

We will analyze the number of releases in the pandemic period and compare it with the number of releases in the same period in previous years.

Analysis of new releases

Now we're going to analyze the number of new releases during the quarantine period related to **Hip-Hop/Rap** genre.

Hip-Hop/Rap releases during quarantine period

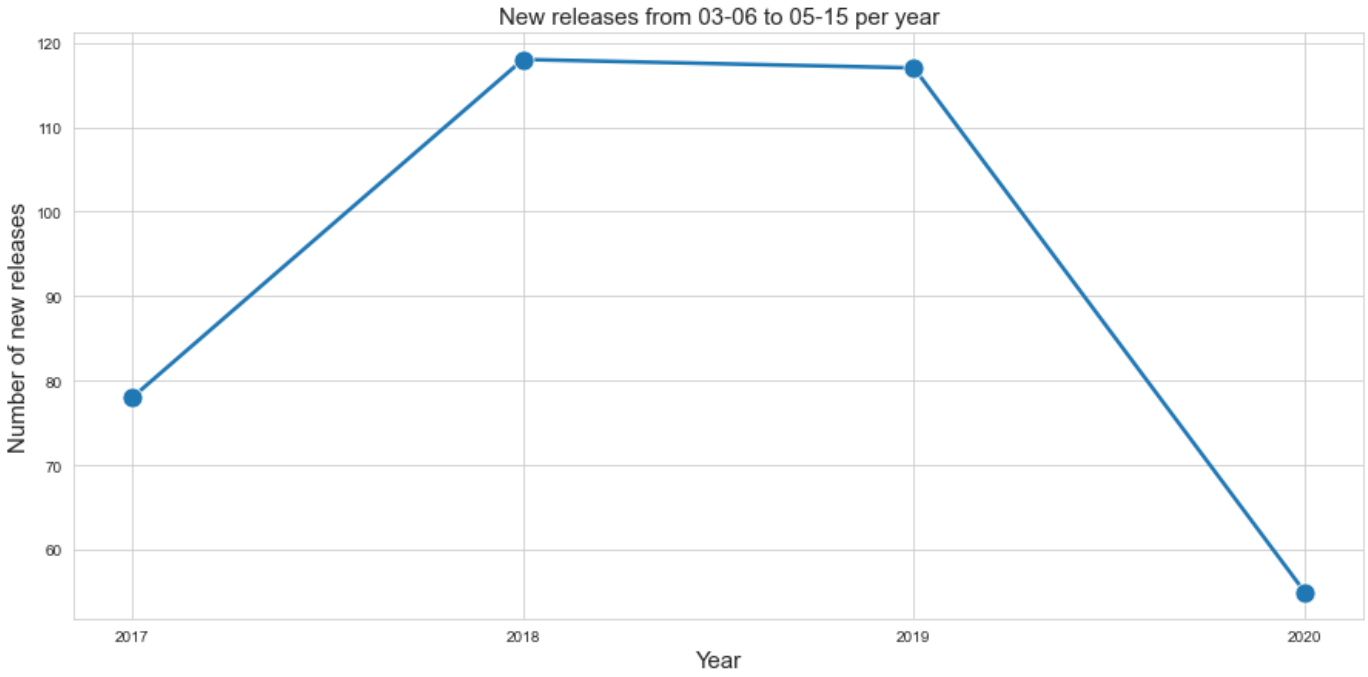
In the following table we have the total number of releases for each year in the period from the **6th of March** to the **15th of May**, which spans the entire lockdown period in Italy. (sources= http://www.salute.gov.it/portale/nuovocoronavirus/dettaglioNotizieNuovoCoronavirus.jsp?lingua=italiano&menu=notizie&p=dalministero&id=4184 and https://www.rainews.it/dl/rainews/articoli/coronavirus-Fase-2-ecco-come-si-riparte-il-18-maggio-6a0c52c3-2f99-4170-95de-e6d221e7ff21.html)

year	number of releases
2017	78
2018	118
2019	117
2020	55

The following plot shows the number of new releases in the period from **03-06** to **05-15** for each year.

As we can see in the plot, we have an increasing trend in the number of releases in the different years which drops in 2020.

```
Text(0.5, 1.0, 'New releases from 03-06 to 05-15 per year')
```



Probably no major italian Hip-Hop/Rap releases

By looking at the plot above, we can notice a drastic decrease in the number of releases in 2020. As discussed previously, most of the significant peaks in the number of streams are related to Italian rappers' influential album releases.

We first build a dataframe which contains the number of new releases per artist in the period from **03-06** to **05-15** for each year. For example we will end up with something like:

```
| Artist | Year | New |
| Queen  | 1972 | 15  |
| Queen  | 1980 | 12  |
| Michael Jackson | 1979 | 10 |
```

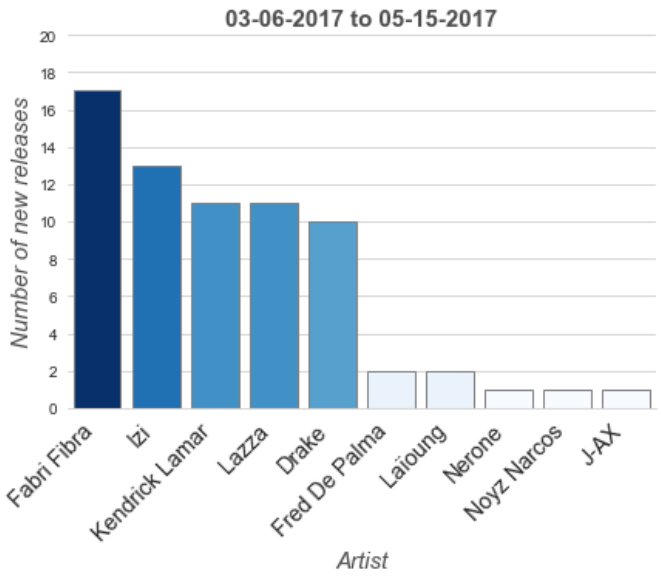
This means that **Queen** released 15 new tracks in the period from **03-06-1972** to **05-15-1972**, they also released 12 new tracks in the period from **03-06-1980** to **05-15-1980** and **Michael Jackson** released 10 new tracks in the period **03-06-1979** to **05-15-1979** and so on.

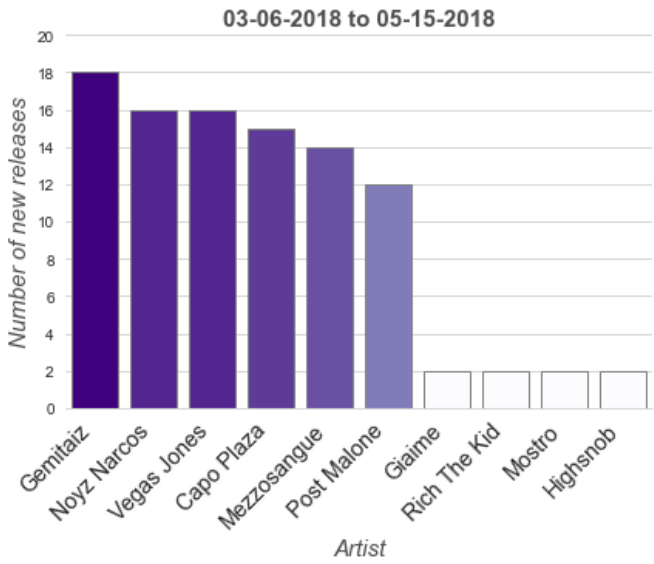
As we can see, the number of Italian releases in the period considered, in 2017, 2018, and 2019 are quite a lot if compared with 2020, which had only a release by ***Nitro*** right at the beginning of the quarantine period: he released his album ***GarbAge*** on the **6th March 2020**.

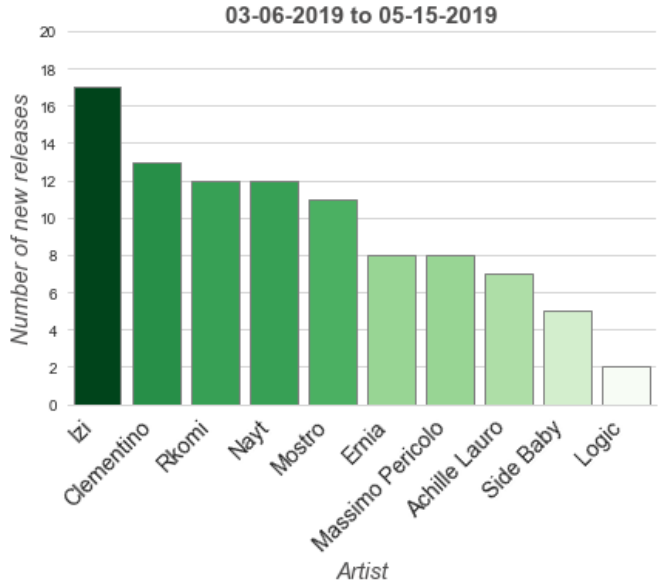
The most acute observer could notice something peculiar: we have a decent number of new releases from ***Ghali*** and ***Marracash*** in the period from **03-06-2020** to **05-15-2020**, but ***Ghali*** released his album ***DNA*** on the **21st of February 2020** while ***Marracash*** released ***Persona*** on the **31st October 2019**.

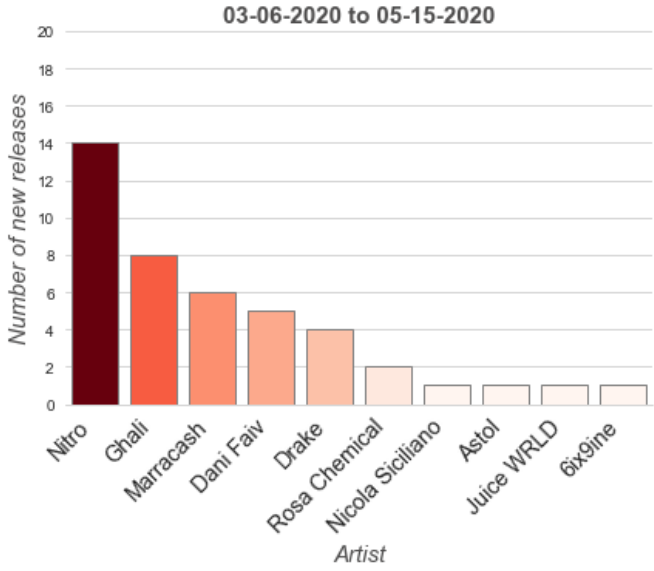
Someone might wonder why there are releases by artists that do not correspond to the release of new albums and are neither new singles. We will analyze this particular situation in the following.

```
[(0.0, 20.0)]
```









Analyzing Marracash anomaly

If we analyze the file **it_2020-03-27--2020-04-03**, we can notice 6 new releases by Marracash!

We retrieve also the top 200 corresponding to the week of the release of the album, which is the **1st November 2019**. In the second dataframe, we can notice 15 tracks from ***Marracash*** in the top 200, which basically are all the tracks from the album ***Persona***.

The tracks are the following:

1. Body Parts - I denti
2. Qualcosa in cui credere - Lo scheletro (feat. Gué Pequeno)
3. Quelli che non pensano - Il cervello (feat. Coez)
4. Appartengo - Il sangue (feat. Massimo Pericolo)
5. Poco di buono - Il fegato
6. Bravi a cadere - I polmoni
7. Non sono Marra - La pelle (feat. Mahmood)
8. Supreme - L'ego (feat. Tha Supreme, Sfera Ebbasta)
9. Sport - I muscoli (feat. Luchè)

10. Da buttare - Il ca**o
11. Crudelia - I nervi
12. G.O.A.T - Il cuore
13. Madame - L'anima (feat. Madame)
14. Tutto questo niente - Gli occhi
15. Greta Thunberg - Lo stomaco (feat. Cosmo)

We can clearly notice that all the tracks contained in the first dataframe, except for ***SPORT + muscoli (RMX)*** are also present in the second dataframe, but they have different release dates and different IDs!

Why? We're going to discover this later.

First thing first, a track which is present in both dataframes and check if it's actually the same doing a query to the Spotify API. We're going to use ***SUPREME - L'Ego***

`2020-03-27`

	Position	Track Name	Artist	ID	release date	new
2	3	SPORT + muscoli (RMX) (feat. Lazza, Paky, Luch...	Marracash	7srDOLMwjMC07Dzs7mqCnv	2020-03-27	1
22	23	BRAVI A CADERE - I polmoni	Marracash	3me7fQdjgO8NHscb3xPaBa	2020-03-27	1
28	29	SUPREME - L'ego (feat. tha Supreme & Sfera Ebb...	Marracash	5EpBEqf5bnKeJ3p4zA1Sod	2020-03-27	1
33	34	NEON - Le Ali (feat. Elisa)	Marracash	76cEI7LczlKZ1yLAHIoPBJ	2020-03-27	1
35	36	CRUDELIA - I nervi	Marracash	5dsz7MTrNdN9aMTrnp7sOG	2020-03-27	1
95	96	MADAME - L'anima (feat. Madame)	Marracash	3go5ZuMoFTKl5ugai8nAKi	2020-03-27	1

`2019-11-01`

	Position	Track Name	Artist	ID	release date	new
0	1	SUPREME - L'ego (feat. tha Supreme & Sfera Ebb...	Marracash	020wG4EHTwBy5h2rbpxqvK	2019-10-31	0
1	2	CRUDELIA - I nervi	Marracash	1OeIJITrlUR4qss2kywMEn	2019-10-31	0
3	4	APPARTENGO - Il sangue (feat. Massimo Pericolo)	Marracash	2tpfOljBdskx8G331TjRps	2019-10-31	0
4	5	QUELLI CHE NON PENSANO - Il cervello (feat. Coez)	Marracash	6l1Jnqfu7GAOV4MDgA2bNP	2019-10-31	0
5	6	QUALCOSA IN CUI CREDERE - Lo scheletro (feat. ...	Marracash	39hRcGieqFaKfZ7LN91Vy5	2019-10-31	0
6	7	MADAME - L'anima (feat. Madame)	Marracash	2wL96n9ToLGhekNTFrPhxj	2019-10-31	0
8	9	BRAVI A CADERE - I polmoni	Marracash	4SpV49wwekArulAVjEez3i	2019-10-31	0

	Position	Track Name	Artist	ID	release date	new
9	10	SPORT - I muscoli (feat. Luchè)	Marracash	3HVm4g9bZHo0aMsjzGS7Ro	2019-10-31	0
10	11	BODY PARTS - I denti	Marracash	4WOXQYfq3fYVRMcaXuHBAD	2019-10-31	0
11	12	NON SONO MARRA - La pelle (feat. Mahmood)	Marracash	0pl6P6vTEB9pNABaD53SgS	2019-10-31	0
13	14	G.O.A.T. - Il cuore	Marracash	50c3vq2uUkSV8QrcbcPWW3	2019-10-31	0
15	16	POCO DI BUONO - Il fegato	Marracash	4ykwdc532KqgY4z8CJcUwO	2019-10-31	0
20	21	DA BUTTARE - Il ca**o	Marracash	2auOwPOpAxhGmJflaxS0kQ	2019-10-31	0
21	22	TUTTO QUESTO NIENTE - Gli occhi	Marracash	5Lwin5QMDUiC5W39iJ6nsi	2019-10-31	0
24	25	GRETA THUNBERG - Lo stomaco (feat. Cosmo)	Marracash	38VKvSaq4ic5k5zCHshjMo	2019-10-31	0

We can notice that the album is the same, and also all the artist infos are exactly the same! But there are 2 major differences:

1. **Release date**: the considered song in the first dataframe has release date **2020-03-27** while in the second dataframe has release date **2019-10-31** which is coherent with the offical release date of the album ***Persona***
2. **Number of tracks**: the considered album has 2 different number of tracks! **17** if we consider the ID song from the first dataframe and **15** if we consider the second dataframe.

From the previous points and dataframes, we can conclude that 2 tracks have been added later which are ***SPORT + muscoli (RMX)*** and ***NEON - Le Ali***, which are present in the first dataframe, but not in the official track list of the album ***Persona***.

We discovered that adding extra tracks such as remixes, collaboration, featurings ecc... to an album, will automatically update the release date of all the tracks in the album!

In fact, if you visit the following link which contains the album ***Persona***: https://open.spotify.com/album/19iZTn6IM82raMquk5Z7Ul where you can see the release date is 2019

While if you visit this link: https://open.spotify.com/album/3ZOt77e63uMgJXU7xcFpqu which contains the same album, but the release date is 2020.

It is worth noticing that the same track from the 2 album versions has the same identical number of streams, this could give some hints about spotify music storage and format. It seems that the track is actually the same, but it has different ID and release date, and it's present in two different versions of the same album.

The track we're using as test is ***SUPREME - L'ego***.

```
-------------First result-------------
Artist infos:
```

```
[{'external_urls': {'spotify': 'https://open.spotify.com/artist/5AZuEF0feCXMkUCwQiQlW7'},
  'href': 'https://api.spotify.com/v1/artists/5AZuEF0feCXMkUCwQiQlW7',
  'id': '5AZuEF0feCXMkUCwQiQlW7',
  'name': 'Marracash',
  'type': 'artist',
  'uri': 'spotify:artist:5AZuEF0feCXMkUCwQiQlW7'}]
'Album name: Persona'
'Release date: 2020-03-27'
'Album number of tracks: 17'
------------Second result-------------
Artist infos:
[{'external_urls': {'spotify': 'https://open.spotify.com/artist/5AZuEF0feCXMkUCwQiQlW7'},
  'href': 'https://api.spotify.com/v1/artists/5AZuEF0feCXMkUCwQiQlW7',
  'id': '5AZuEF0feCXMkUCwQiQlW7',
  'name': 'Marracash',
  'type': 'artist',
  'uri': 'spotify:artist:5AZuEF0feCXMkUCwQiQlW7'}]
'Album name: Persona'
'Release date: 2019-10-31'
'Album number of tracks: 15'
```

Going back to COVID releases analysis

The previous considerations hold also for ***Ghali***. He added two new songs to his ***DNA*** album released on the **21st February 2020**, which are ***Cacao*** and ***Hasta la vista***.

Everything we said so far gives us two important hints:

1. We noticed that new releases inside an already existing album refresh the release date of the whole album (and also the tracks).
2. Our estimate of the number of releases per period is in general an upward estimate of the real number of releases

Said that, ***Ghali*** and ***Marracash*** released only 2 and 3 new tracks in that period, which are remixes and/or featuring.

The only major album release by an italian artist have been ***GarbAge*** by ***Nitro***. We can conclude that the pandemic had a relevant impact on the work of artists, as we had already thought, and the lack of new releases impacted on the average number of streams on Spotify.

Remarks: Our conclusion are based on the top 200 weekly, thus we can't conclude that the whole Spotify platform had a lower amount of active users/streams, and also that the overall number of releases by artists has decreased. We can only conclude that in the mainstream/commercial scene, there haven't been new releases by famous artists.
It could be interesting to filter and analyze only italian releases and compare their impact on the Spotify charts with respect to international releases, but unfortunately Spotify doesn't give any information about the language or release nation of the songs.

In could be also interesting to analyze and further investigate if there have been significant changes overall on the platform, if there have been a major change in the listened genres and so on but, again, we don't have access to the needed data.

The following article shows how people were more keen to listen to lo-fi chill music during the quarantine if compared to the "normality": https://blog.chartmetric.com/covid-19-effect-on-the-global-music-business-part-1-genre/

Unfortunately we can't show/prove these changes since our data is limited to the top 200, mainstream music hasn't seen major changes in the most listened genres.

Comparing streams trend for different nations

In the following cells we're going to plot the streams trend for different nations worldwide. We decided to restrict our analysis only to 6 nations due to the high amount of time needed to download the data from the internet (due to rate download limitations of the provided APIs).

In particular we focused on major ***EU countries + USA and Brazil***, focusing on countries that faced lockdown measures against COVID.

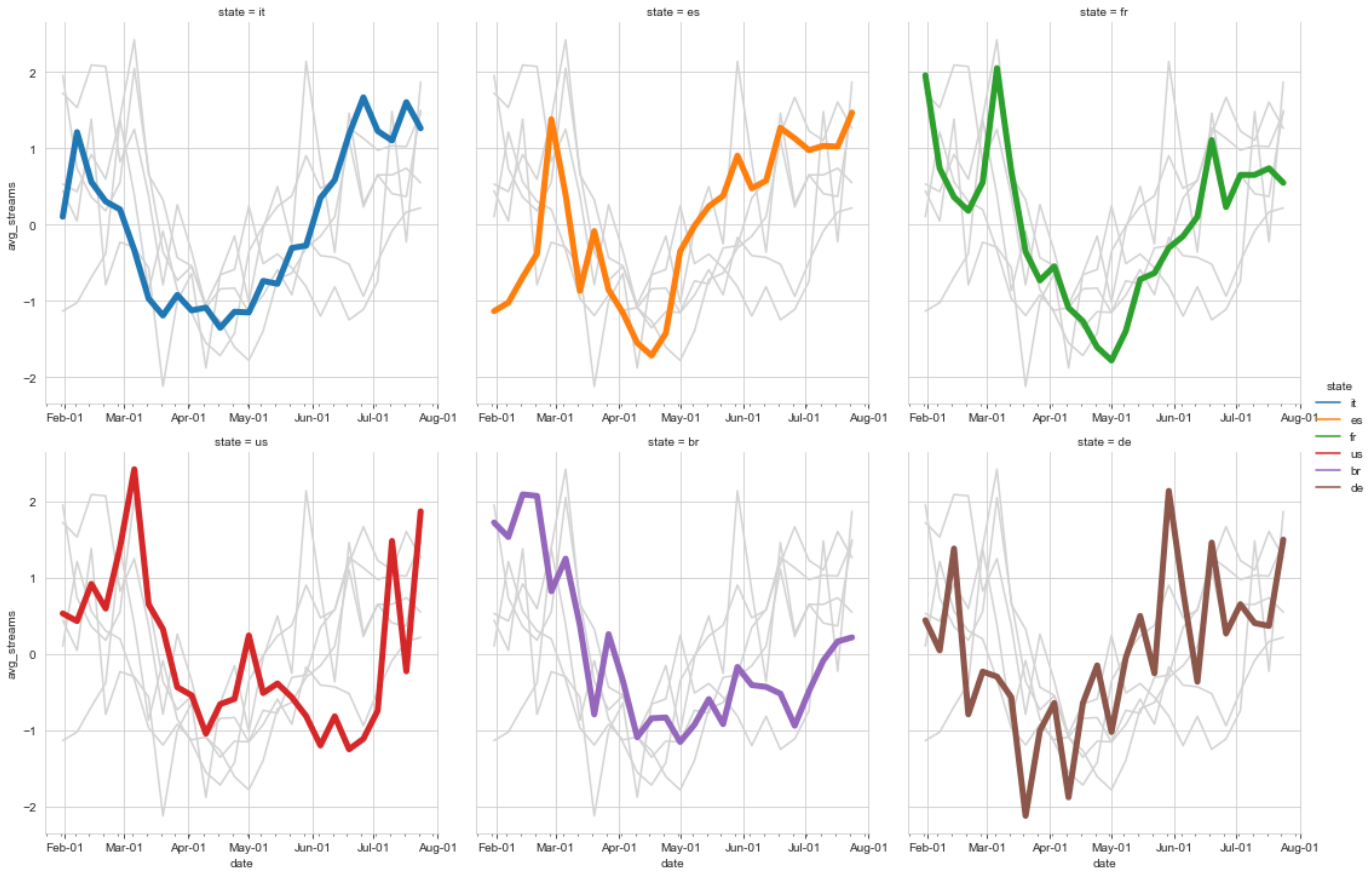
Data Normalization

We applied a Z-Score normalization to our data for each nation. Data among nations was heterogeneous, compare streams on an absolute level generated untreatable plots: USA had an average number of streams 10 times higher than other countries.

Therefore we normalize our data to make it comparable.

Stream trend for each nation.

The plots show a common decreasing trend around 15 weeks among all states with an increasing trend towards summer.



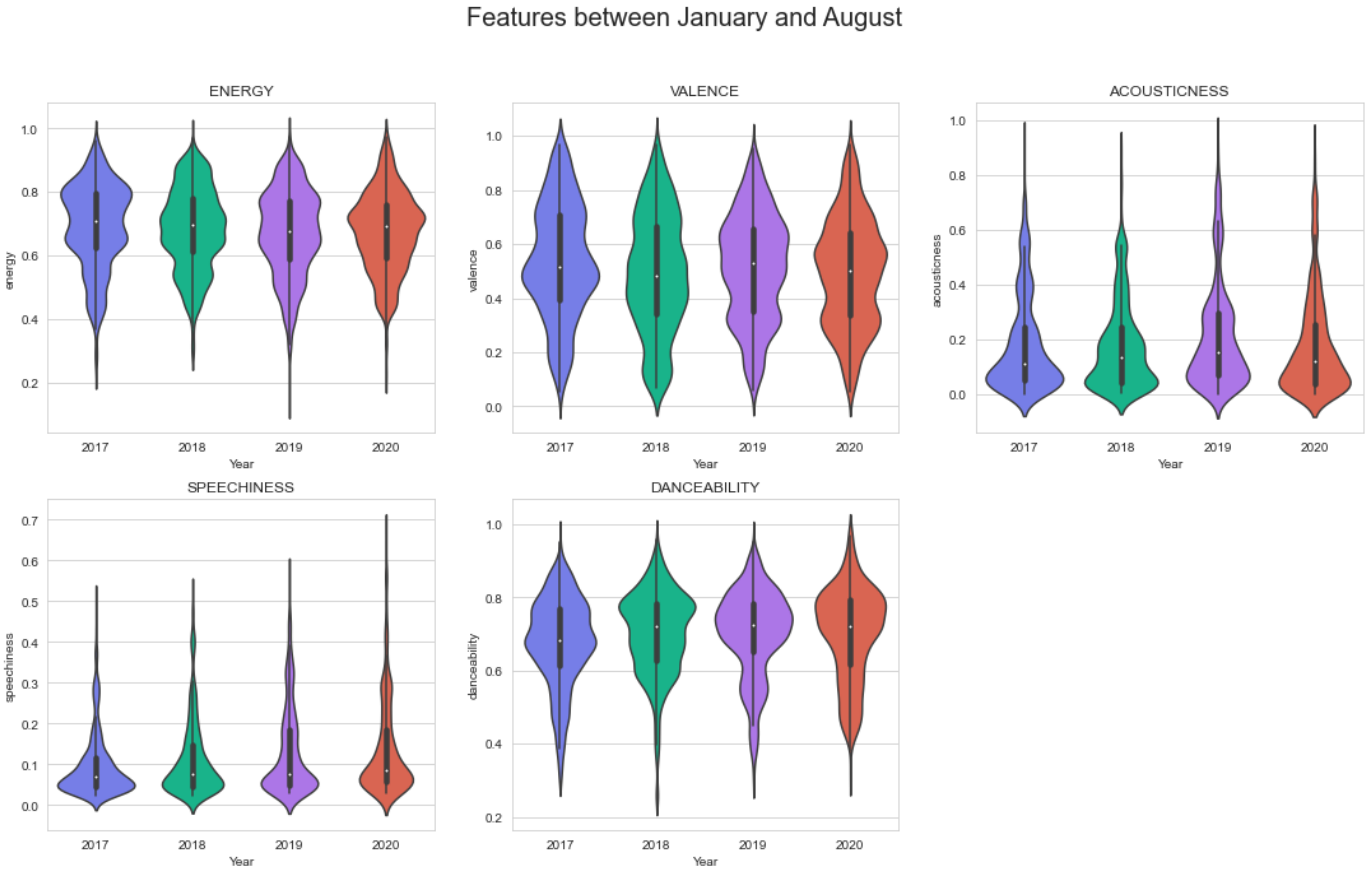
Features Analysis in Italy

The following charts show Spotify's musical features over time and the pandemic's influence on them. As previously mentioned, those results refer only to the top 50 because many song remained in the top 200 for many weeks, so that no absolute conclusions concerning global music trends can be drawn. However, we could still infer some critical information on the most popular songs and see if there has been any shift in the users' preferences in the period under study.

Features distribution

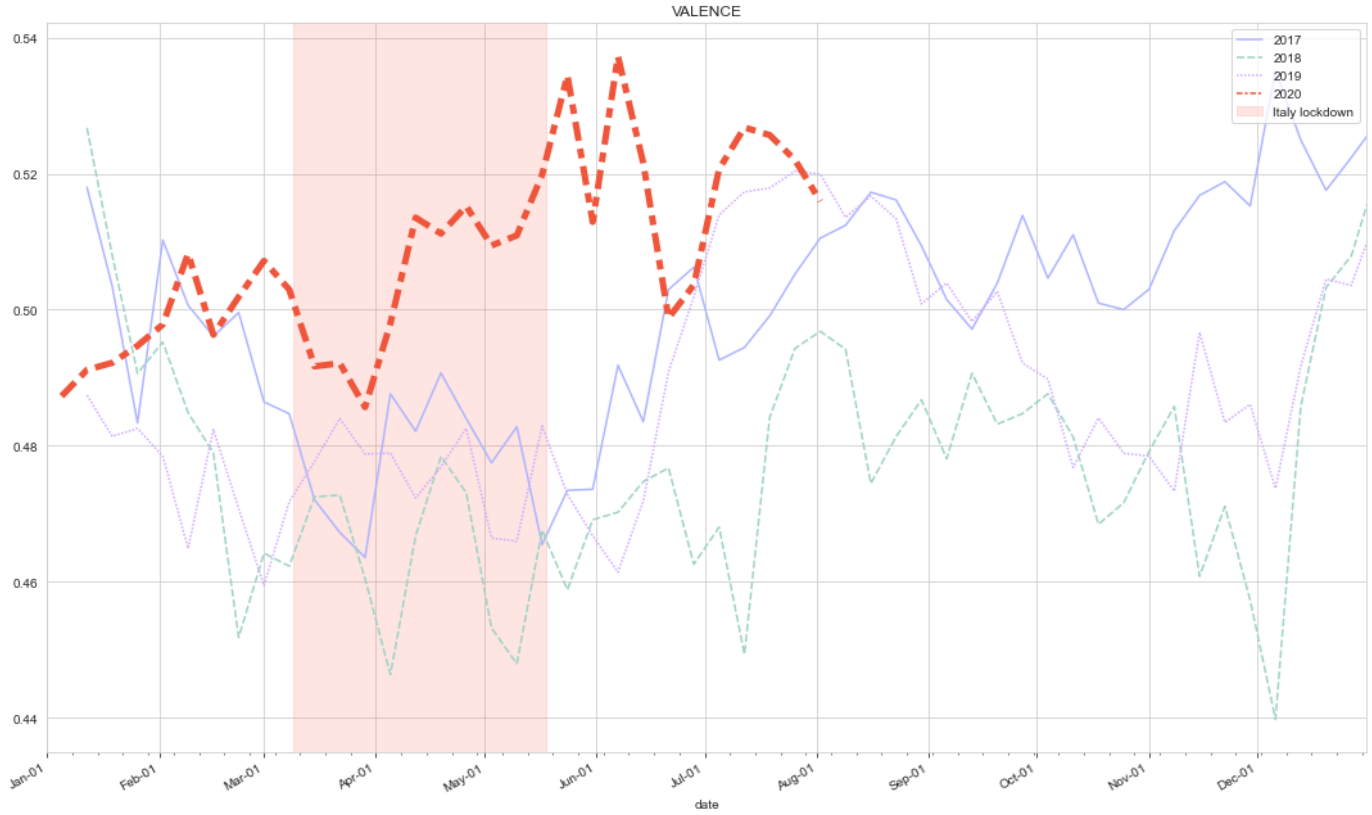
Features distributions don't change significantly across years.
The only relevant change is in speechiness: it has a higher maximum value in 2020 if compared to previous years.

```
Text(0.5, 0, 'Year')
```



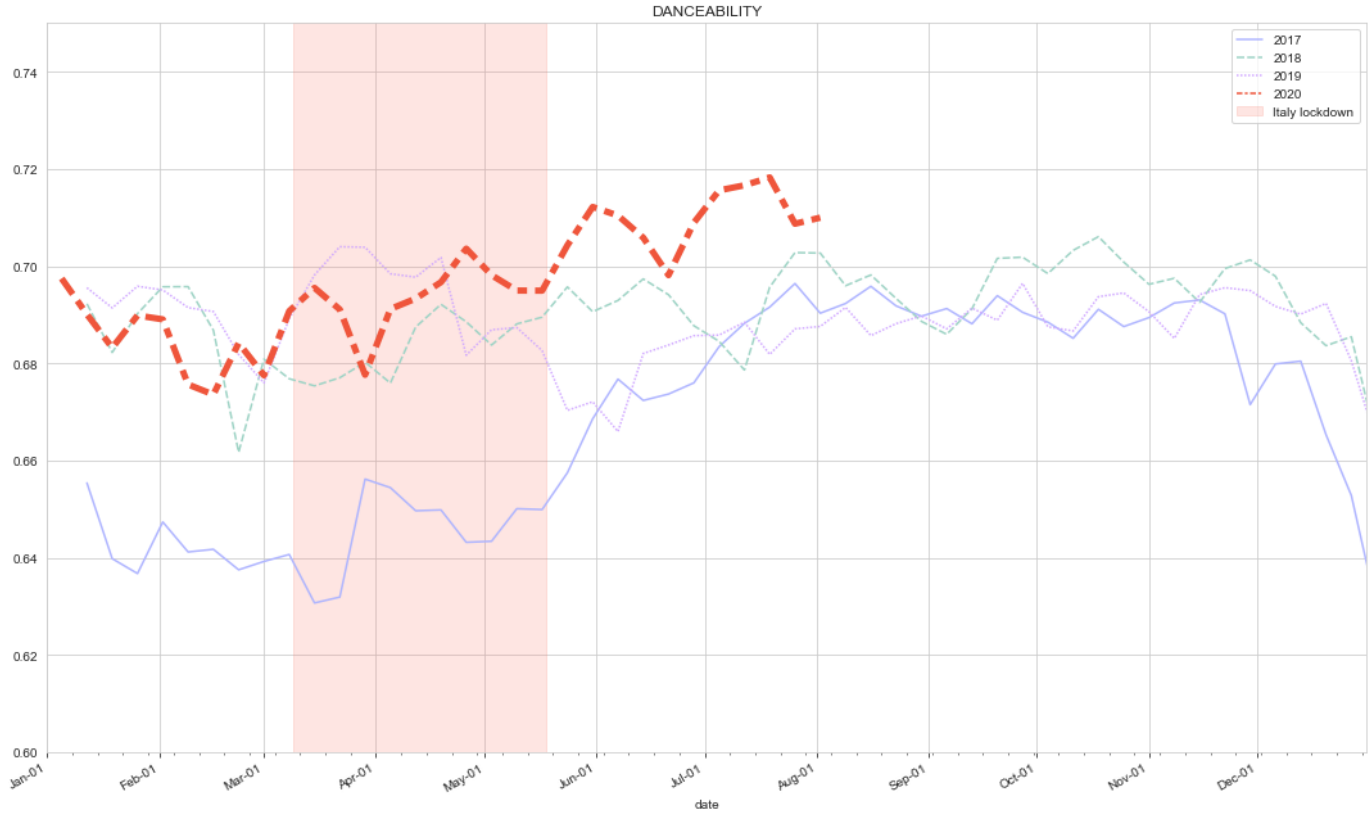
Valence

As defined by Spotify, Valence *measures the musical positiveness conveyed by a track, using a score between 0.0 and 1.0.*
This parameter undergoes some fluctuations over the year and shows some periodicity. In particular, one can notice that every year the average valence increases around Christmas time. The steep increase during the lockdown period seems to contrast the periodicity and could be attributed to the tendency to look for comforting and cheerful music in that period of social isolation. However, considering the significant variance of this feature, combined with the stability of the top 50 in terms of musical genres, may indicate that this abrupt growth happened by chance.



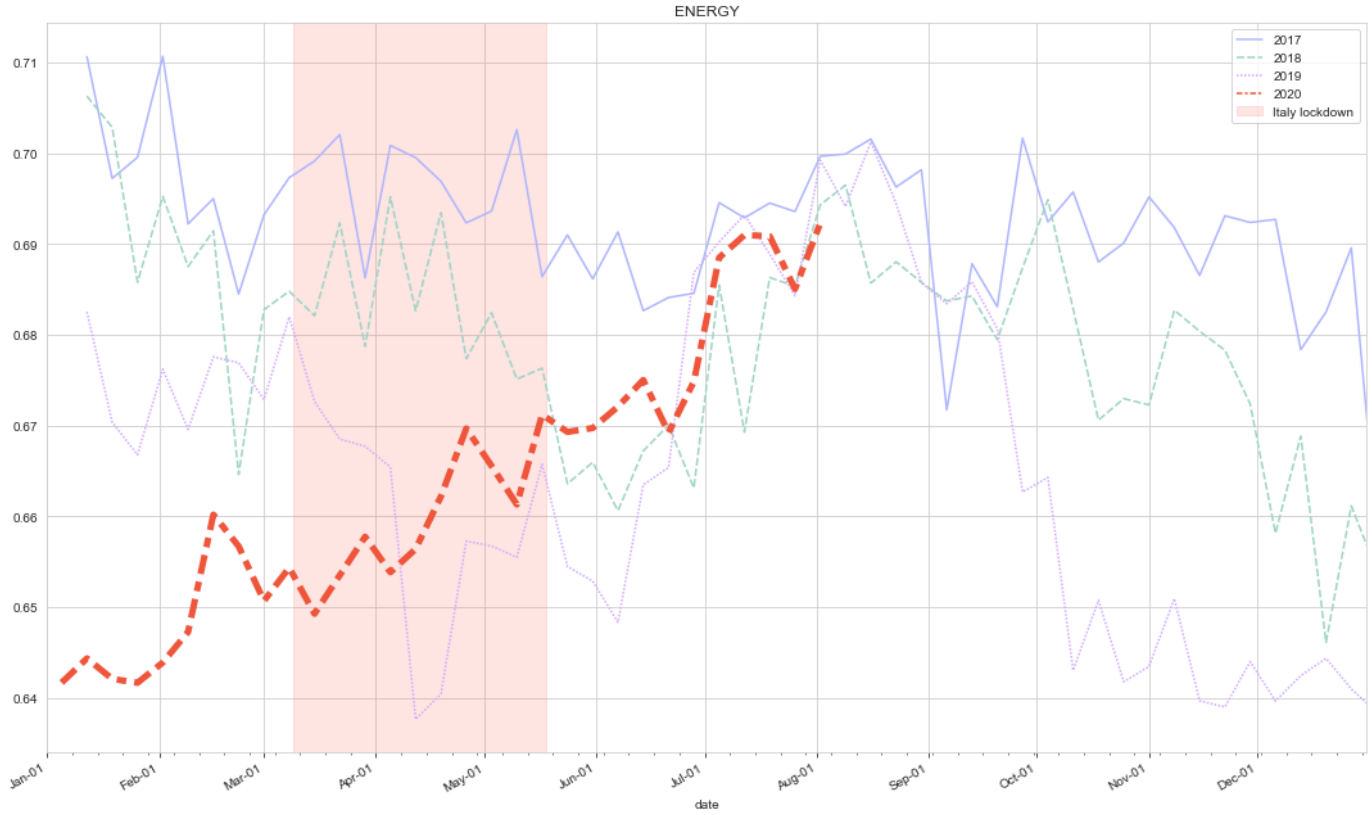
Danceability

Danceability *describes how suitable a track is for dancing based on a combination of musical elements, including tempo, rhythm, and overall regularity.* Danceability increased in 2018 and since then has been almost constant, without showing particularly relevant variations even during the lockdown period.



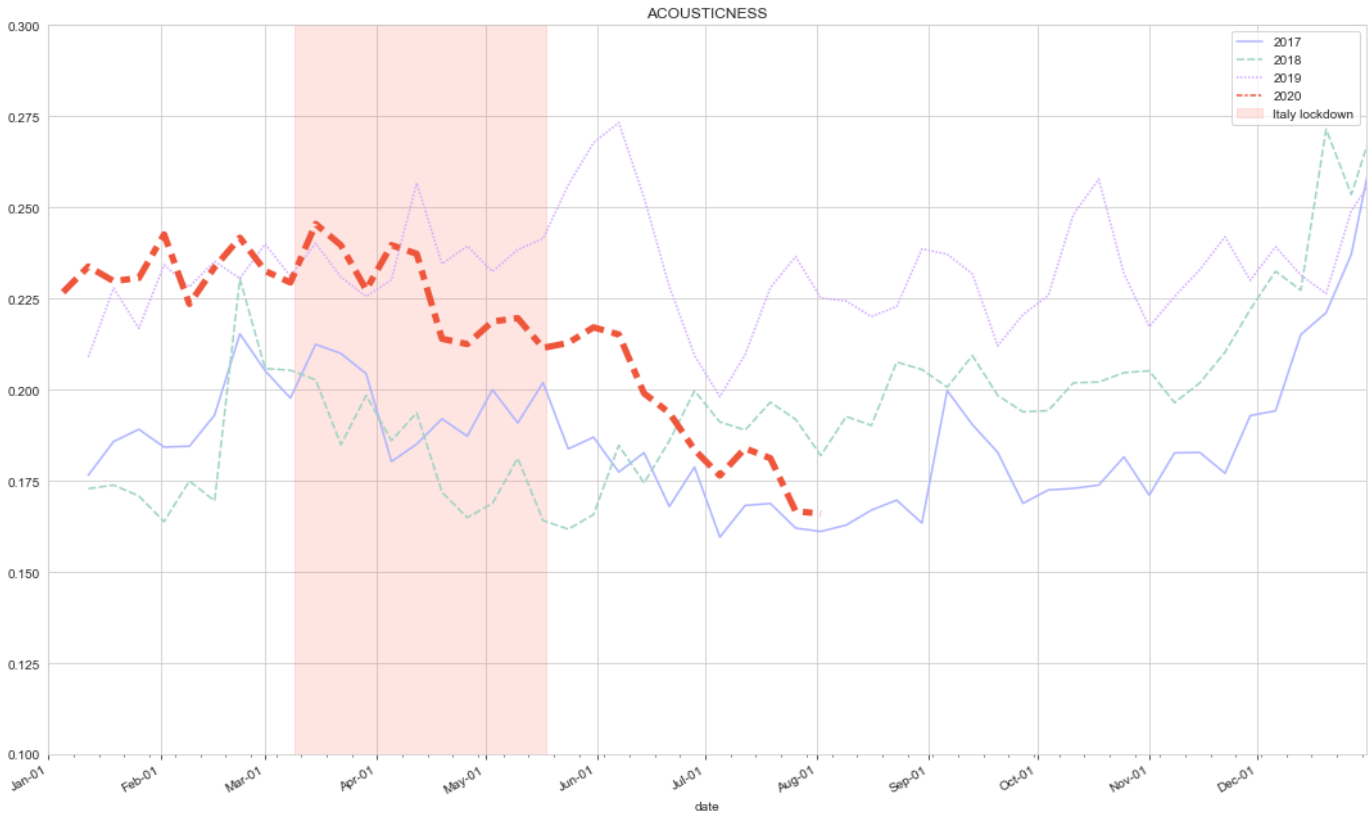
Energy

Energy *measures the overall intensity and activity of a track, as a function of its speed, loudness, and noisiness.*



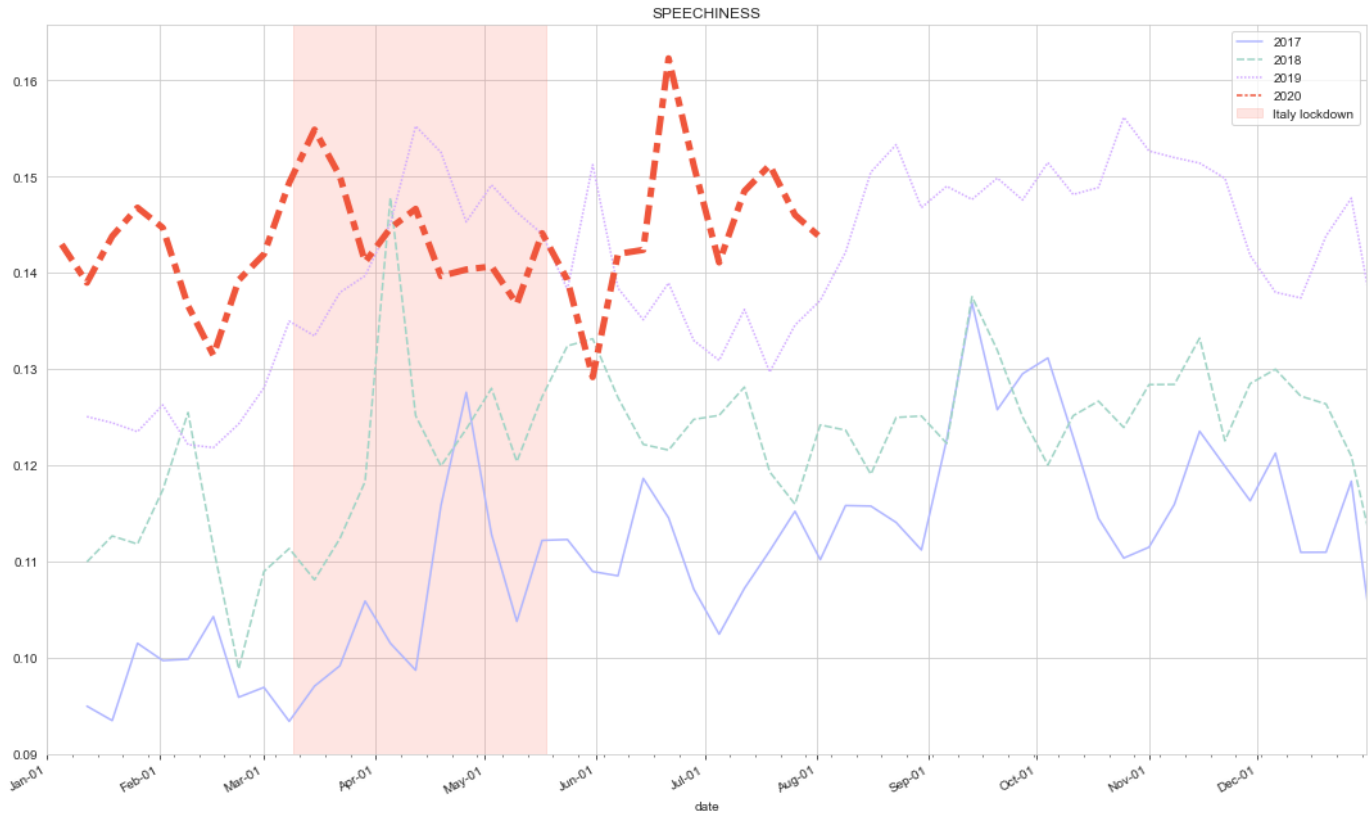
Acousticness

Acousticness, as the name suggests, *measures the presence in a track of acoustic intstruments.* It is possible to observe a significant drop during the lockdown period. Nevertheless, considering the complete histoty of this feature, this trend shows some periodicity, indicating that it may not be due to the pandemic.



Speechiness

Speechiness measures the presence of spoken words in a track. The plot shows no particular COVID-related trends, but a steady increase over the past four years, indicating the growing popularity of rap/hip-hop music, that dominates the top 200, as shown previously.



Radar Chart

This radar chart is intended to show the comparison between features in different periods of

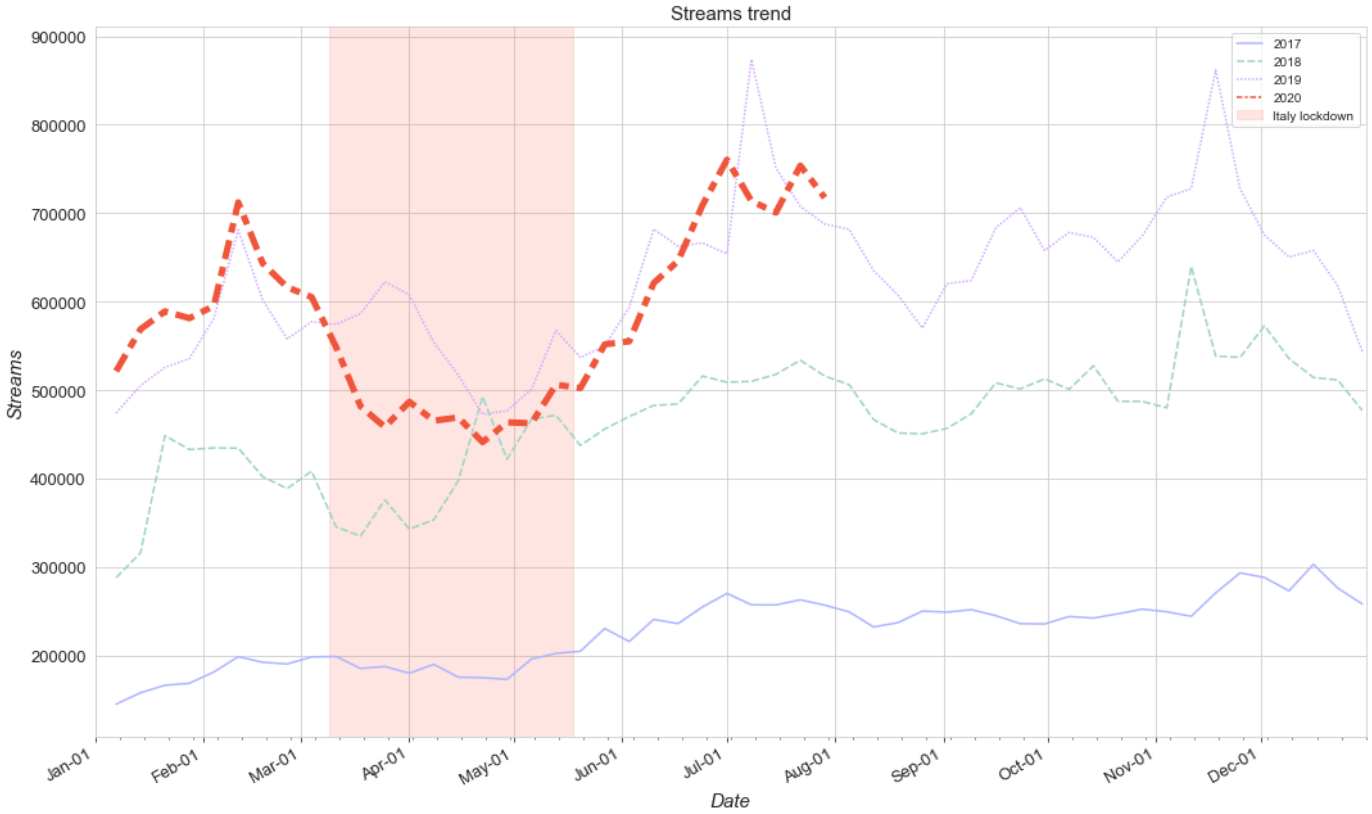
time. It's an interactive plot to allow the user to compare and see the difference of all the features in many period of time.

Date 1: 2017-1

Date 2: 2017-1

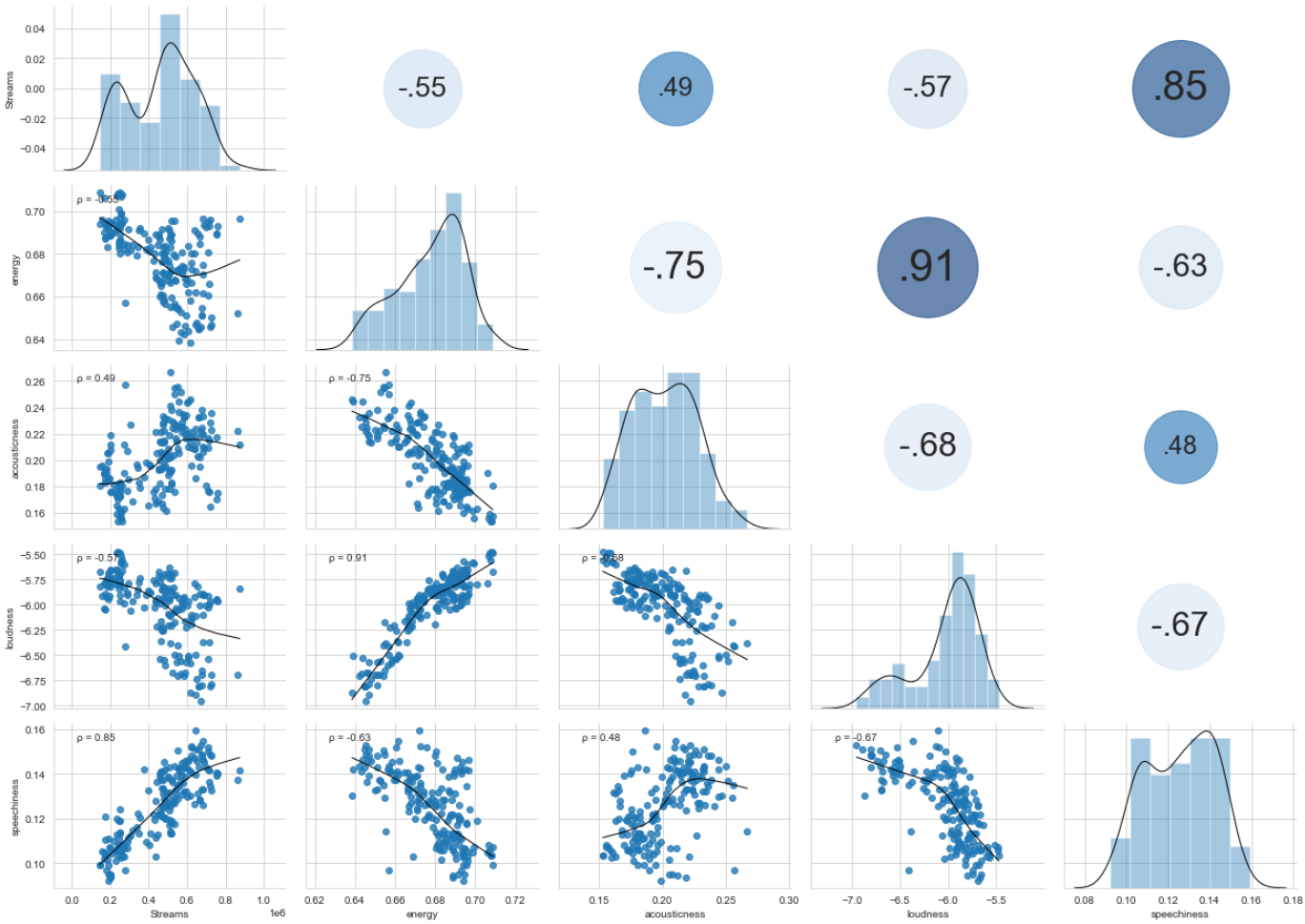
Conclusions

Our analysis showed that, limited to the data available to us, COVID 19 **could** have had an impact on music streamings.
From the very first plot we noticed a decrease in the number of streams during the quarantine period in Italy.

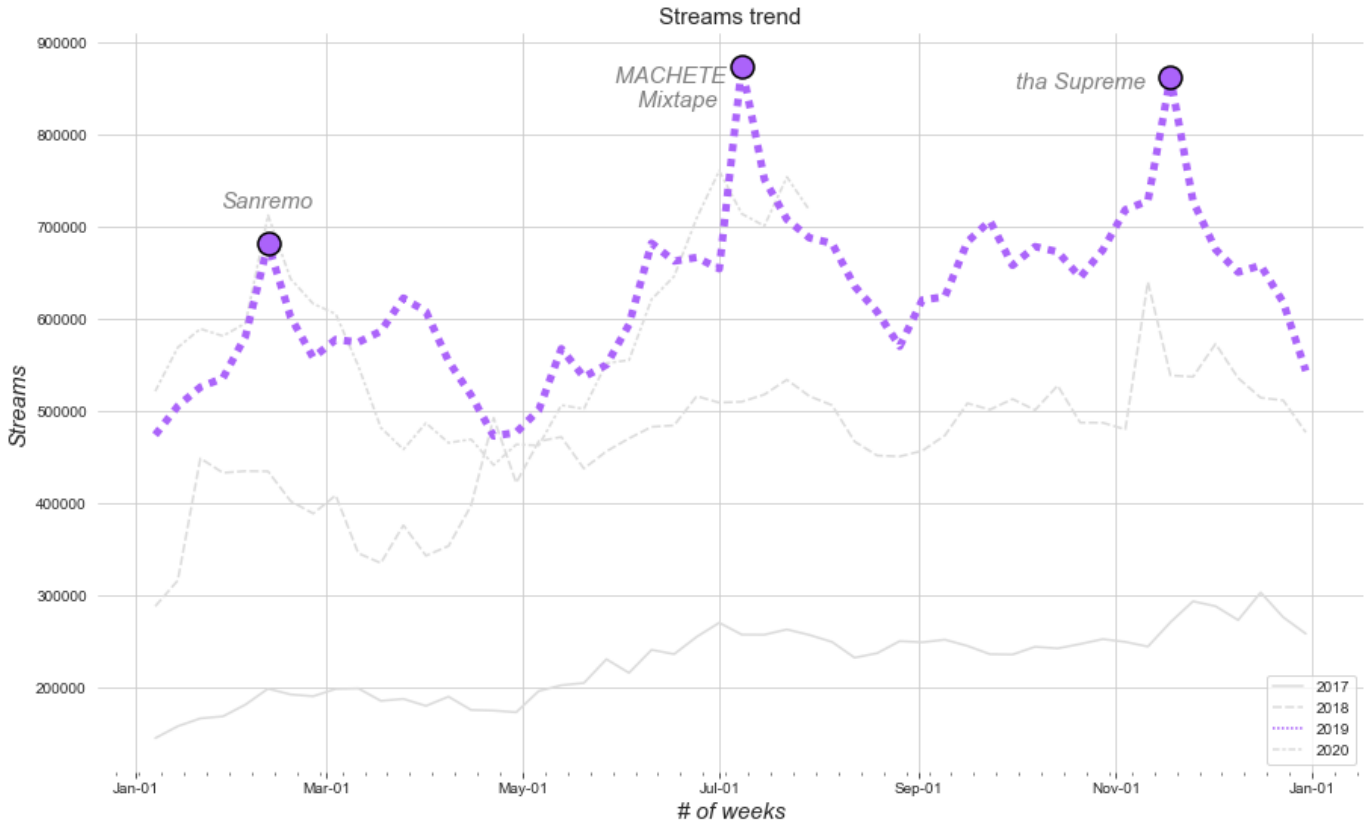


We also noticed a strong correlation between the number of streams and the `speechiness` feature, thanks to our domain knowledge (being italian and musicians) we supposed that Hip-Hop, Rap and Trap song had a major impact on Spotify streams trend.

In order to confirm this hypothesis, we decided to analyze the major peaks in the streams trend plot. This analysis confirmed our hypothesis: major peaks in streams corresponded to musical events (**Sanremo festival**) and **Rap, Hip-Hop and Trap releases**.




```
<matplotlib.legend.Legend at 0x293515b9580>
```



As a consequence, given that the decreasing number of streams could be caused of a smaller number of Rap, Hip-Hop and Trap releases, we filtered our dataset selecting only the top200

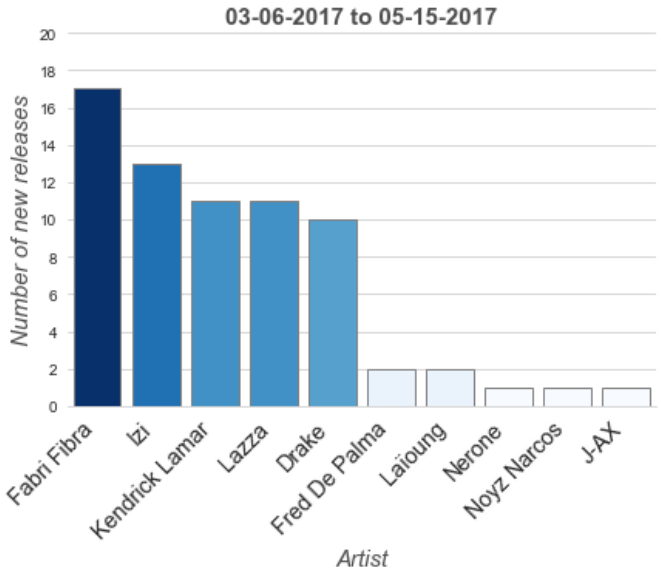
corresponding to the quarantine period (from March to May) for all years and we plotted the number of new Hip-Hop/Rap releases.

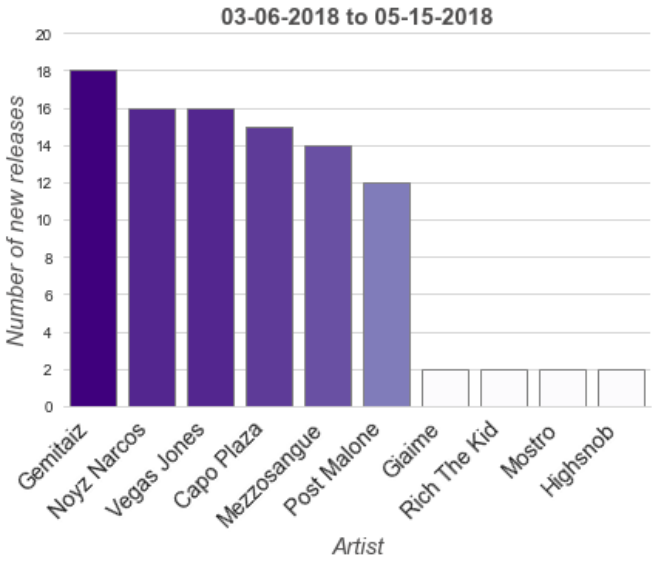
It resulted that in 2020 only **Nitro** released a new album which **didn't impact positively the streams trend**: we suppose that this has been caused by the coincidence of the release date of the album with the lockdown announcement by the italian government, with people's attention being caught by the news.

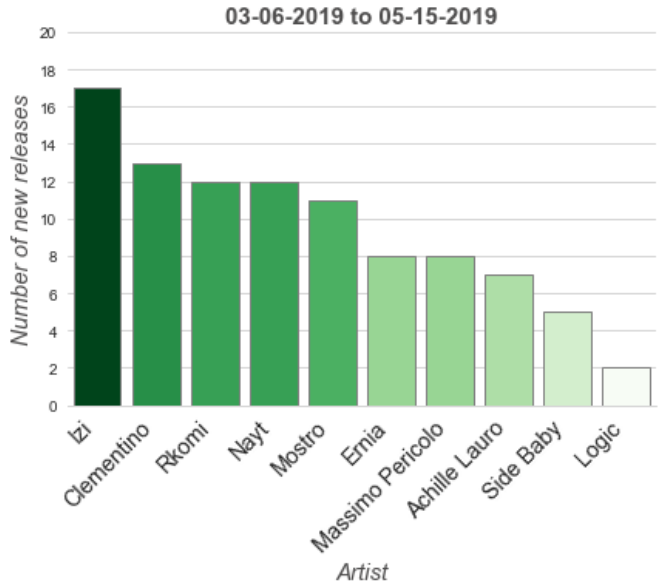
As we already discussed in our report, new releases by **Ghali** and **Marracash** are actually **fake data**: the addition of late remixes to the album, refreshes the release date of all the songs contained inside it.
The most famous songs of the album (which usually remain for a while in the top 200) are considered again as new releases, even if they're not.

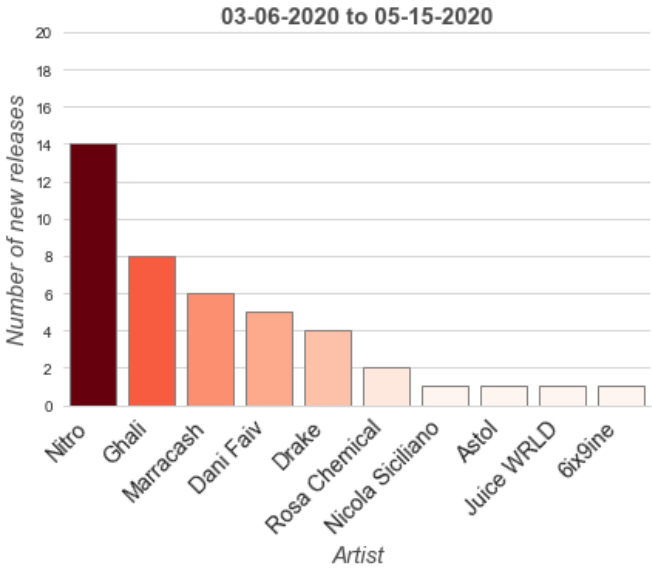
Thus our hypothesis is not so wrong, probably the negative trend of streamings has been caused by a lack of new releases.

```
[(0.0, 20.0)]
```









We also noticed no major changes in audio features, their distribution remained the same. This is kind of natural: mainstream music is dominated by Pop, Hip-Hop and Rap songs, thus

analyzing only the top 200 will not show significant changes in features distribution.

However, our conclusion are limited only to the top200 and are not valid in general.