



Optimization of Spotify User Recommendations for a Better User Experience Based on the Effects of Global Catastrophic Events

***Correlations between music listening trends and
hardship levels from COVID***

Our project:

Explore how **COVID affected user** preferences:

- Are they listening to **happy music** to cheer up?
- Or are we listening to **sad music** to cope?
- Does this **vary by the country**?

2019 was a (relatively) normal year. Were music trends different in 2020?



Business Impact:

Spotify can use this data to improve their algorithms and offer more personalized suggestions to users in different regions, that could potentially provide a better experience to the user.

Music and Happiness

- Most of the existing music recommendation systems use collaborative or content based recommendation engines. [1] They don't consider emotion.
- Research has found that music may induce or potentially regulate positive and negative emotions. [2]
- Approximately 96% of people listened to music in order to regulate their emotions. [3]

Spotify: Business success through personalization

- Users are more likely to interact and retain the service if they get better playlist personalization
- Users that listen to recommended playlists, 81% are premium users. [4]

Questions we wanted to answer?

Did the pandemic affect **how people** (globally and locally) **listen to music**?

Is there a **correlation** between COVID-19 severity and music preferences by country?



Methodology:

Data: Spotipy, Spotify API and Our World in Data - COVID-19.

*CSV files for COVID data and Spotify Top Charts (global and per country), Spotify API **for key metrics to determine trends.***

Key metrics: Valence, COVID cases and COVID deaths.

- **Valence.**
As defined by Spotify: it describes the musical positiveness conveyed by a track.
- *COVID cases and deaths.*

Number of cases and deaths during a timeframe.

1

Convert input CSVs into DataFrames

COVID data in csv files from <https://ourworldindata.org> [5] and Spotify Top Charts at spotifycharts.com. They were read and transformed into DataFrames using Pandas.

2

Obtain track features from API

Using the Spotify API and the track names from the Spotify DataFrame, we got track information regarding Valence

3

Delete faulty/ repeated data

Some tracks were not found, and some information was not available in certain countries. These values in the DataFrame were dropped.

4

Plot key indexes over time

Using the dates in which the data was obtained, we merged the DataFrames and plotted valence.

5

Obtain correlations between variables

Correlations between variables were obtained by using scatter plots and linear regression analysis.

Position	Track Name	Artist	Streams	URL	Year	Month	Track ID	Danceability	Valence	Energy
1	7 rings	Ariana Grande	71467874	https://open.spotify.com/track/4msK75pk3pA33pzPVNtBE	2019	1	6ocbgoVGwYJhOv1Ggl9NsF	0.778	0.317	0.327
2	Sunflower - Spider-Man: Into the Spider-Verse	Post Malone	5093889	https://open.spotify.com/track/5093889	2019	1	3KkXRkHbMCARz0aVIEt68P	0.76	0.479	0.913
3	Wow.	Post Malone	31165644	https://open.spotify.com/track/6MWtB6iixylwun0YzU6DEP	2019	1	7xQAFvXzm3AkraOtGPWIZg	0.829	0.539	0.388
4	thank u, next	Ariana Grande	27672341	https://open.spotify.com/track/2rPE9A1vEgShuXxzR2tZH	2019	1	3e9HZxeyfWwjeYpAMmWSSQ	0.717	0.653	0.412
5	Without Me	Drake	27576694	https://open.spotify.com/track/5pTujcrUXASCNwRaWNHR1C	2019	1	7GSMWtB6iixylwun0YzU6DEP	0.908	0.669	0.662
6	Calma - Remix	Pedro Capó	23772814	https://open.spotify.com/track/5iww1NieX7WWjnCgY5TH4	2019	1	5wz21NnWwJlCgY5TH4	0.826	0.773	0.761
7	Dancing With A Stranger (with Normani)	Sam Smith	22238899	https://open.spotify.com/track/6Qs4SXO9dwPj5GKvVOv8Ki	2019	1	6Qs4SXO9dwPj5GKvVOv8Ki	0.741	0.52	0.347
8	Spotify CSV:	Ava Max	21711506	https://open.spotify.com/track/25sgk305KZfy	2019	1	7DnAm9FOTWE3cUvso43HhI	0.72	0.706	0.62
9	Spotify API:	DJ Snake	21435790	https://open.spotify.com/track/4w8niZpiMy6q	2019	1	4w8niZpiMy6qz1mntFA5uM	0.842	0.801	0.617
10	ML	Bad Bunny	19597964	https://open.spotify.com/track/116H0KvKj2Z4RPuVBruDO	2019	1	116H0KvKj2Z4RPuVBruDO	0.817	0.539	0.158
11	Hi	Panic! At The Disco	18891709	https://open.spotify.com/track/1rqQCsm0Qe4l9rUvWncaom	2019	1	1rqQCsm0Qe4l9rUvWncaom	0.579	0.904	0.681
12	Ha	Marshmello	18135105	https://open.spotify.com/track/2dpaYNEQHfRxtZbfNse99	2019	1	2dpaYNEQHfRxtZbfNse99	0.687	0.792	0.671
13	Ad	Paulo Londra	17547601	https://open.spotify.com/track/6FyRXC8tJUh863JcKyWqtK	2019	1	6FyRXC8tJUh863JcKyWqtK	0.767	0.709	0.72
14	SICKO MODE	Travis Scott	17280171	https://open.spotify.com/track/2xLMiifQCjDGFmkHkpNLD9h	2019	1	2xLMiifQCjDGFmkHkpNLD9h	0.834	0.73	0.446
15	Our World in Data CSV:		17174274	https://open.spotify.com/track/2fRZnDFmiqMuO	2019	1	2fRZnDFmiqMuO	0.889	0.496	0.544
16			16829254	https://open.spotify.com/track/2VxeLyX666F8uXCJ0dZF8B	2019	1	2VxeLyX666F8uXCJ0dZF8B	0.572	0.385	0.323
17	Lost in the Fire	Gesaffelstein	15898519	https://open.spotify.com/track/7wFybC8jBH3	2019	1	7wFybC8jBH3	0.658	0.671	0.166
18	Se	Anuel AA	15649853	https://open.spotify.com/track/5W83ErFkO3a	2019	1	5W83ErFkO3a	0.807	0.803	0.706
19	ZE	Jak Black	15564793	https://open.spotify.com/track/0FZ4Dmg8JJAPJ	2019	1	0FZ4Dmg8JJAPJ	0.861	0.603	0.504
20	Sal	anald	14560398	https://open.spotify.com/track/5kfNlritmkNE8mU	2019	1	5kfNlritmkNE8mU	0.551	0.44	0.341
21	Eat	benny blanco	14028996	https://open.spotify.com/track/7FEG80cy8juXBCD2nrqdW	2019	1	7FEG80cy8juXBCD2nrqdW	0.56	0.68	0.319
22	Drip Too Hard (Lil Babv & Gunna)	Lil Babv	13585875	https://open.spotify.com/track/78QR3Wp35dqA	2019	1	78QR3Wp35dqA	0.397	0.662	0.389
23	Each CSV was done both with		3523043	https://open.spotify.com/track/5JEx7HbmVHQQ	2019	1	5JEx7HbmVHQQ	0.575	0.758	0.493
24	global and individual country		13473222	https://open.spotify.com/track/3OI2xnObFdKV9pmRD2t9x8	2019	1	3OI2xnObFdKV9pmRD2t9x8	0.791	0.587	0.536
25	data.		13432564	https://open.spotify.com/track/5itQtNxoWxtJmI	2019	1	5itQtNxoWxtJmI	0.887	0.625	0.606
26		Sheck Wes	13264212	https://open.spotify.com/track/1xzBco0xcoJEDXxtI7Jxrr	2019	1	1xzBco0xcoJEDXxtI7Jxrr	0.729	0.625	0.261
27	returning breaks Like a Heart (feat. Miley Cyrus)	Mark Ronson	12917943	https://open.spotify.com/track/27rdGxbavYJeBphck5MZAf	2019	1	27rdGxbavYJeBphck5MZAf	0.601	0.794	0.244
28	when the party's over	Billie Eilish	12787942	https://open.spotify.com/track/14JzyD6FIBD5zQwV5P0ZYI	2019	1	14JzyD6FIBD5zQwV5P0ZYI	0.367	0.111	0.198

Data Transformation

Exploration

Analysis

Spotify CSV:

- Track name
- Streams
- Artist

Our World in Data CSV:

- COVID cases
- New COVID cases
- COVID deaths

Each CSV was done both with global and individual country data.

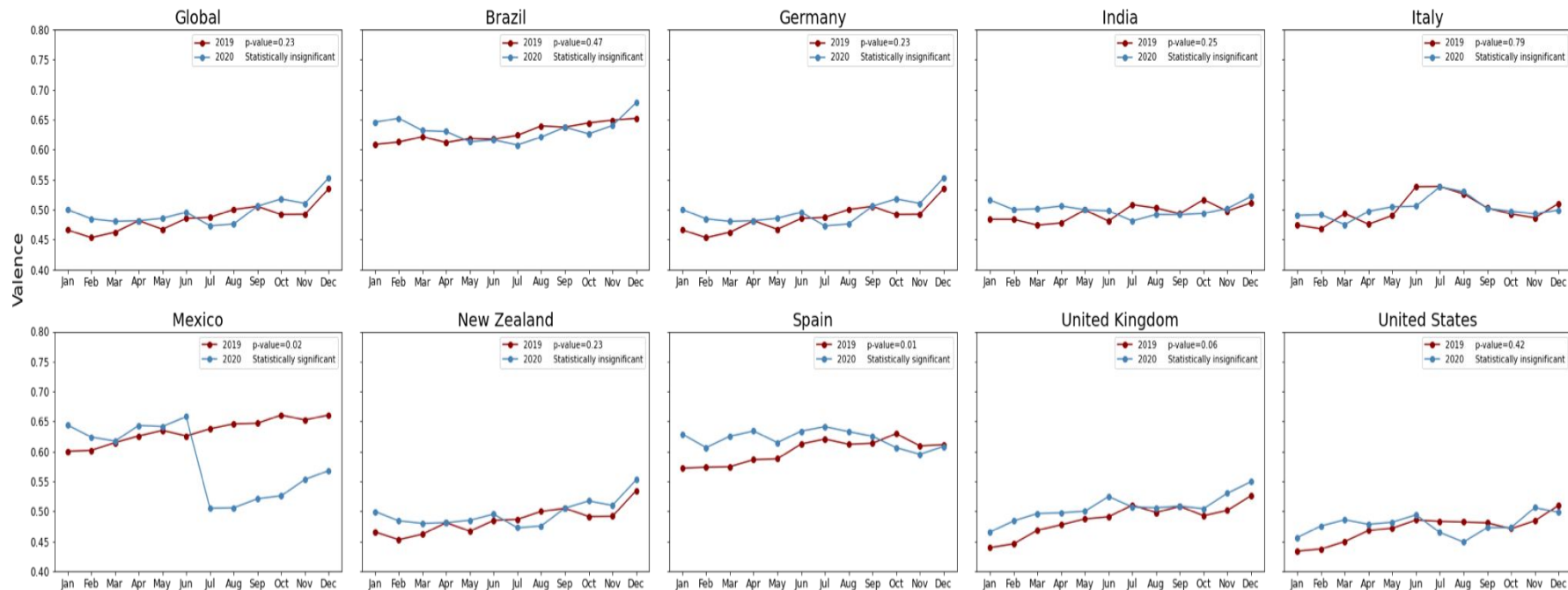
Spotify API:

- Track ID
- Month
- Year
- Audio features

Pandas Analysis:

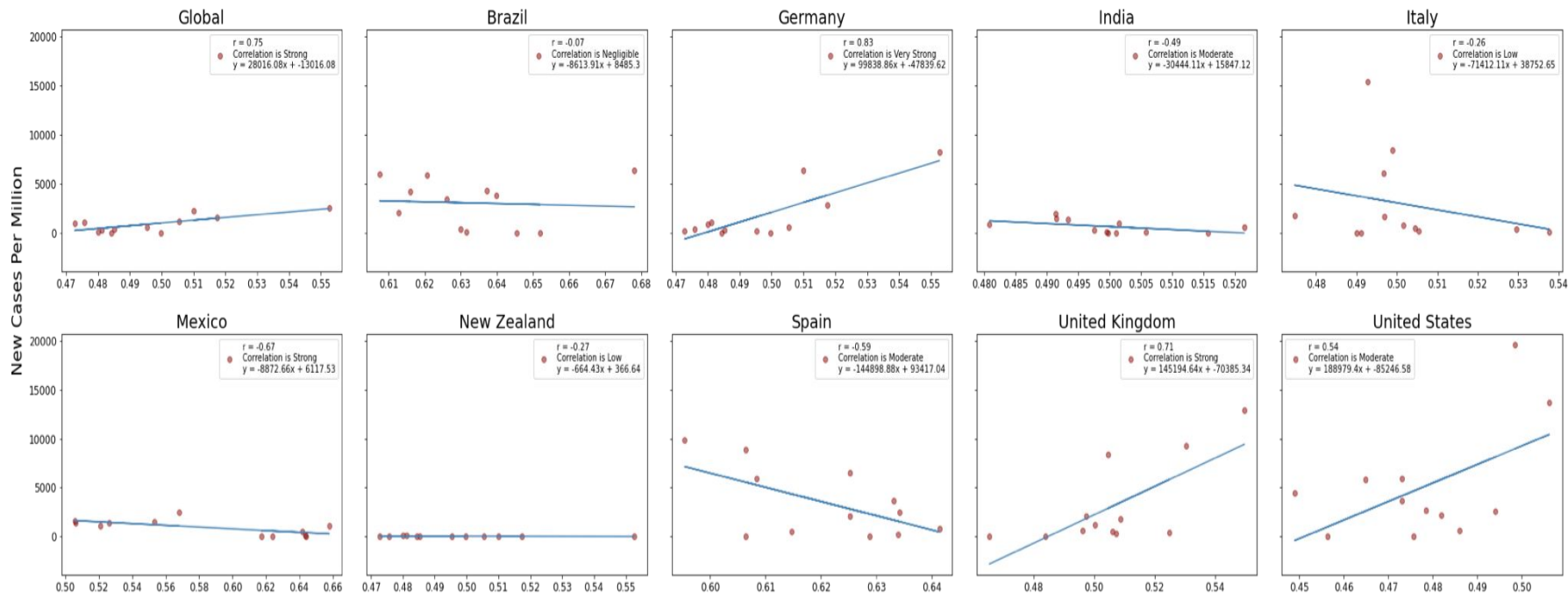
- Mean values for audio features.
- COVID cases and deaths per million.
- Time analysis and correlation.

SPOTIFY LISTENER VALENCE PREFERENCE: 2019 VERSUS 2020



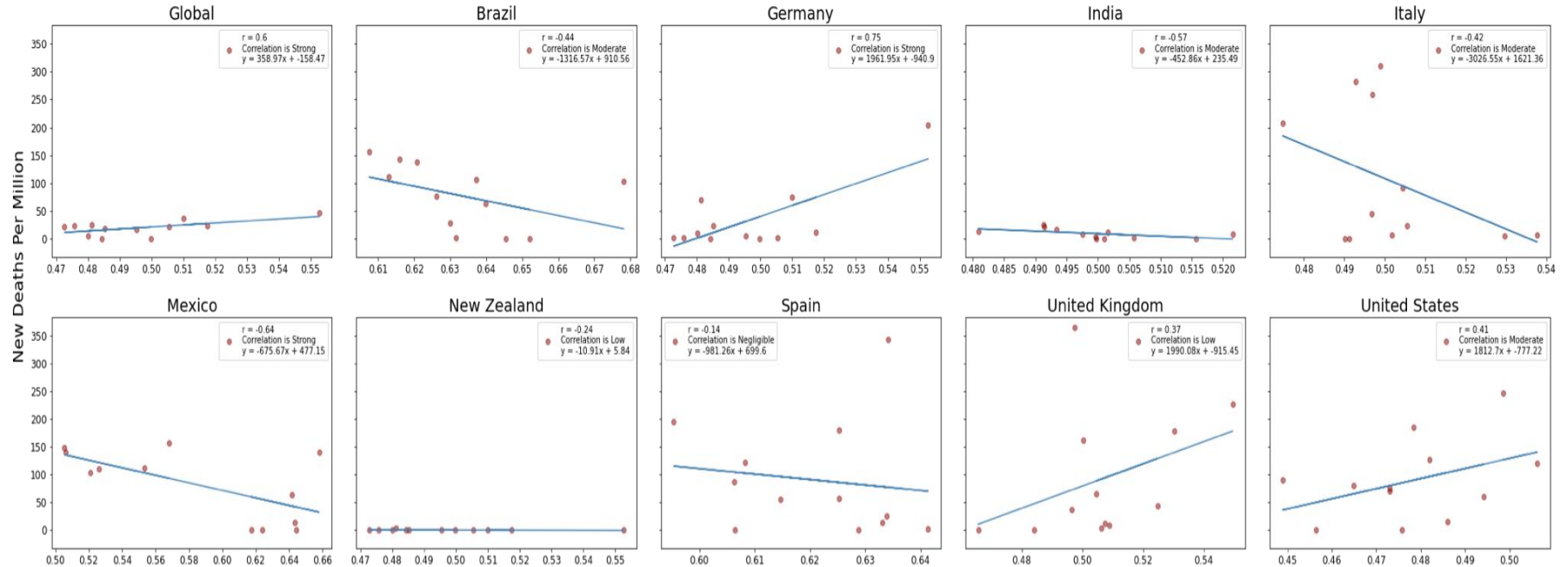
Results: Song Valence Over Time

SPOTIFY LISTENER VALENCE PREFERENCE VERSUS NEW COVID CASES



Results: Correlation Valence vs. New Cases per Million

SPOTIFY LISTENER VALENCE PREFERENCE VERSUS NEW DEATHS PER MILLION



Results: Correlation Valence vs. Deaths per Million

Results

- Valence preference varies between countries.
 - Hispanic and Latin countries have higher mean valence.
- Mean valence between 2019 and 2020 was significantly different in some countries.
 - México decreased in 2020.
 - Spain increased in 2020.
- Globally, there was no significant change in mean valence.

Valence vs. Cases per Million

- Global tendency shows a strong positive correlation.
- Germany and the US match the trend.
- México, Spain and India show from a moderate to a strong negative correlation.

Valence vs. Deaths per Million

- Global tendency shows a strong positive correlation.
- Germany and the US match the trend.
- México, Brazil, Italy and India show from a moderate to a strong negative correlation.

Can Spotify benefit from these results?

Yes, Spotify can use the data obtained to offer more personalized suggestions to users depending on the country, and optimize exclusive contracts with certain artists that match the 'mood' of each country.

Was there a correlation between COVID-19 effects and the type of music we listen?

Yes. The majority of the countries analyzed presented a correlation.

Are users listening to happy or sad music?

Depends on the country. We can hypothesize that cultural differences and net impact of the pandemic account for this differences, but more study is needed.

Limitations

- Other external events were not accounted for.
 - Elections and political events from around the world.
- Spotify does not have data for key countries, such as China, where the pandemic started.
- Only valence was considered as a key indicator.

What's next?

- Analyzing podcast data.
- Obtain data from other sources, such as Apple Music API.
- Effect on artists and contracts.
 - Did more music get released?
- Analyze other audio features as key indicators.
 - Energy? Danceability? Tone? Tempo?

Q&A Session

References

- [1] Ayata, D., Yaslan, Y., & Kamasak, M. E. (2018). Emotion based music recommendation system using wearable physiological sensors. IEEE transactions on consumer electronics, 64(2), 196-203.
- [2] Cook, T., Roy, A. R., & Welker, K. M. (2019). Music as an emotion regulation strategy: An examination of genres of music and their roles in emotion regulation. Psychology of Music, 47(1), 144-154.
- [3] Lonsdale, A. J., & North, A. C. (2011). Why do we listen to music? A uses and gratifications analysis. British journal of psychology, 102(1), 108-134.
- [4] Sanchez, J. (2018). ALGORITHMS AND CURATED PLAYLIST EFFECT ON MUSIC STREAMING SATISFACTION.
- [5] Oxford University. Our World Data. <https://ourworldindata.org>