

SPOTIFY ON PANDEMIC

CHANGE IN TRENDS OF MUSIC DURING COVID-19 SITUATION

Team 7

20194202 - Zidong Yang

20170850 - Mohamed Almaazmi

20160843 - Peerapon Akkapusit

20160847 - Wasachon Chaisirirat

01 INTRODUCTION

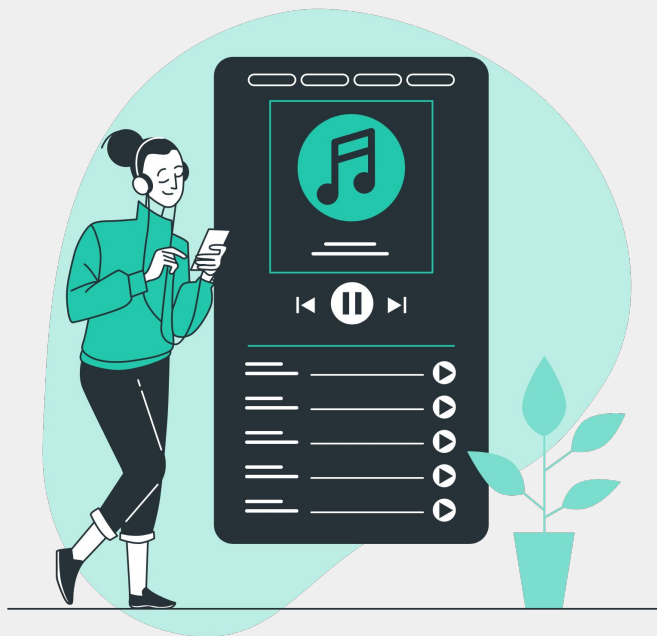
02 DATASETS &
METHODOLOGY

03 RESULTS

04 LIMITATIONS &
ACKNOWLEDGEMENT

05 SUMMARY

INTRODUCTION



WHAT'S HAPPENING WITH MUSIC?

No live concerts, no street shows, and people are turning to streaming services

WHAT MOTIVATES THIS PROJECT?

Music can be a good prediction for other functions such as productivity and happiness so knowing if COVID has affected it is an essential first step.

WHAT DO WE WANT TO DO?

We want to statistically test hypotheses about changes in COVID trends and visualize the factors affecting it.

WHY SPOTIFY?

One of the most popular streaming services and a great indication of music trends as a streaming platform

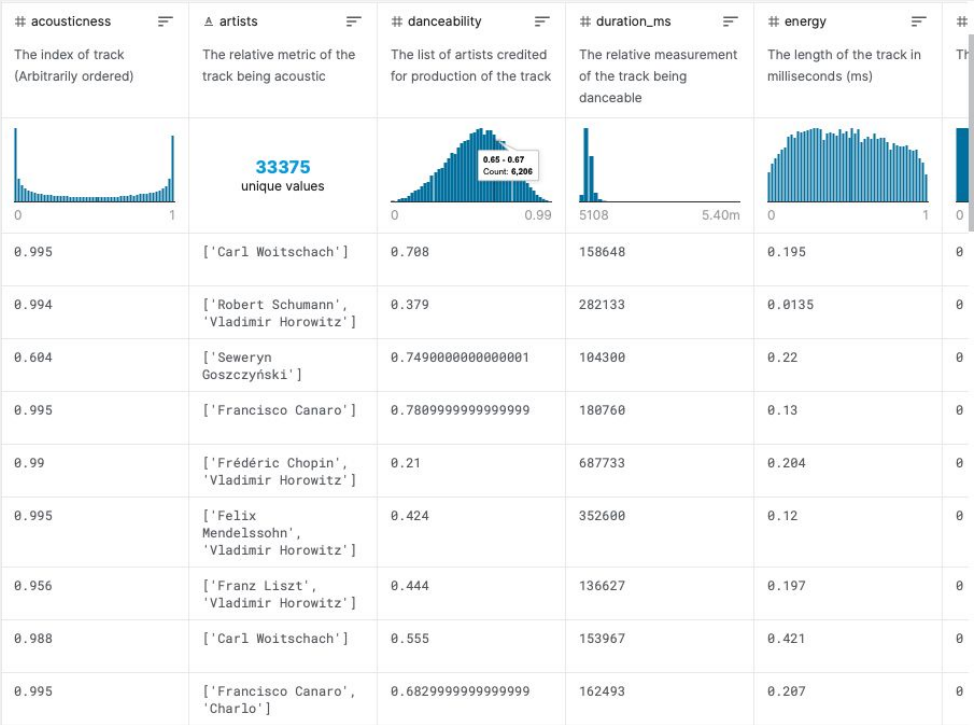


DATASET & METHODOLOGY

DATASET

SPOTIFY DATASET 1921-2020, 160K+ TRACKS

- Artists
- Duration (ms)
- Characteristics of Music
 - Acousticness
 - Danceability
 - Liveness
 - Energy
 - Loudness
 - Speechiness
 - Valence
 - Tempo

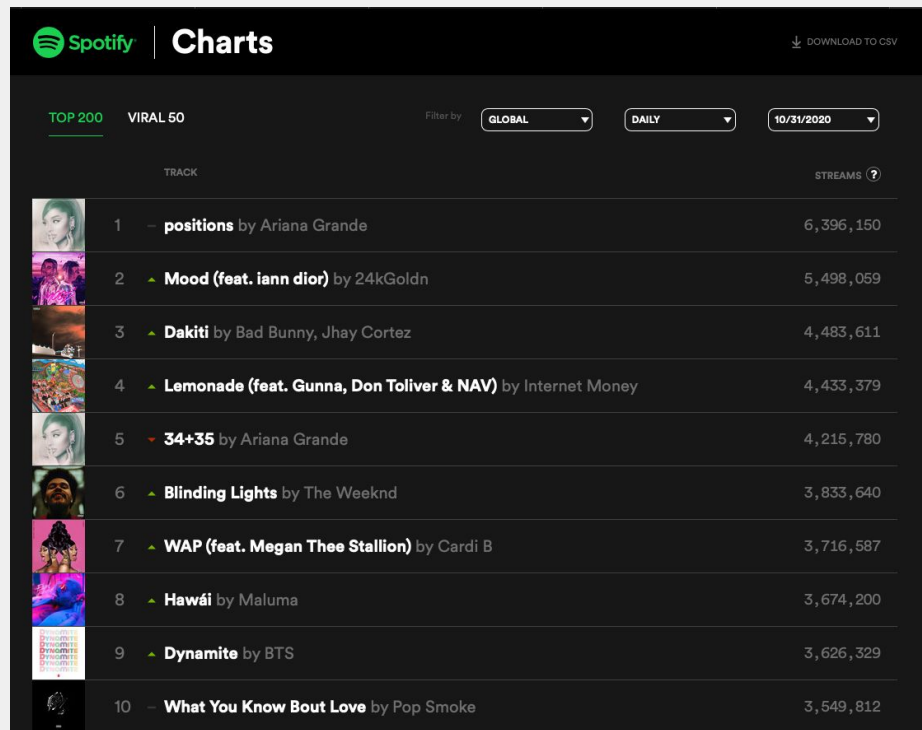












SPOTIFY'S WORLDWIDE DAILY SONG RANKING

- Position on Charts
- Tracks
- # Streams

Custom crawler (written in python) was used to obtain the data from this website using the spotify API

Weekly data from 2019/11/8 to 2020/11/13



Spotify		Charts		↓ DOWNLOAD TO CSV	
TOP 200		VIRAL 50		Filter by	
				GLOBAL	
				DAILY	
				10/31/2020	
TRACK				STREAMS ?	
	1 — positions by Ariana Grande			6,396,150	
	2 ▲ Mood (feat. iann dior) by 24kGoldn			5,498,059	
	3 ▲ Dakiti by Bad Bunny, Jhay Cortez			4,483,611	
	4 ▲ Lemonade (feat. Gunna, Don Toliver & NAV) by Internet Money			4,433,379	
	5 ▼ 34+35 by Ariana Grande			4,215,780	
	6 ▲ Blinding Lights by The Weeknd			3,833,640	
	7 ▲ WAP (feat. Megan Thee Stallion) by Cardi B			3,716,587	
	8 ▲ Hawái by Maluma			3,674,200	
	9 ▲ Dynamite by BTS			3,626,329	
	10 — What You Know Bout Love by Pop Smoke			3,549,812	

DATASET

WORLD-O-METERS COVID CASE NUMBERS

- Country specific and global cases
- Total cases
- New cases

<div> <div>Now</div> <div>Yesterday</div> <div>2 Days Ago</div> <div>Columns</div> <div>Search:</div> </div>													
<div> <div>All</div> <div>Europe</div> <div>North America</div> <div>Asia</div> <div>South America</div> <div>Africa</div> <div>Oceania</div> </div>													
Country, #	Other	Total Cases	New Cases	Total Deaths	New Deaths	Total Recovered	Active Cases	Serious, Critical	Tot Cases/ 1M pop	Deaths/ 1M pop	Total Tests	Tests/ 1M pop	Population
World		64,816,439	+628,028	1,498,299	+12,307	44,921,026	18,397,114	106,457	8,315	192.2			
1	USA	14,309,299	+198,785	279,846	+2,811	8,459,009	5,570,444	26,189	43,123	843	199,864,129	602,322	331,822,744
2	India	9,533,471	+33,761	138,657	+498	8,970,104	424,710	8,944	6,880	100	142,445,949	102,796	1,385,715,753
3	Brazil	6,436,650	+48,124	174,531	+669	5,698,353	563,766	8,318	30,191	819	25,700,000	120,546	213,197,088
4	Russia	2,347,401	+25,345	41,053	+589	1,830,349	475,999	2,300	16,082	281	77,200,071	528,909	145,961,030
5	France	2,244,635	+14,064	53,816	+310	165,563	2,025,256	3,488	34,356	824	20,678,186	316,496	65,334,738
6	Spain	1,682,533	+9,331	45,784	+273	N/A	N/A	2,485	35,980	979	22,992,742	491,692	46,762,467
7	UK	1,659,256	+16,170	59,699	+648	N/A	N/A	1,351	24,388	877	44,059,851	647,589	68,036,766
8	Italy	1,641,610	+20,709	57,045	+684	823,335	761,230	3,616	27,168	944	22,334,342	369,627	60,424,017
9	Argentina	1,440,103	+7,533	39,156	+228	1,268,358	132,589	3,983	31,741	863	3,975,760	87,629	45,370,463
10	Colombia	1,334,089	+9,297	37,117	+183	1,225,635	71,337	2,376	26,102	726	6,535,154	127,862	51,111,013
11	Mexico	1,122,362	+8,819	106,765	+825	829,817	185,780	3,335	8,667	824	2,892,449	22,335	129,502,622

METHODOLOGY

01

COVID DATA VISUALIZATION

Find “cutting points” – points in time where covid starts rising in a country

02

PRINCIPAL COMPONENT ANALYSIS

Determine most relevant features to simplify data

03

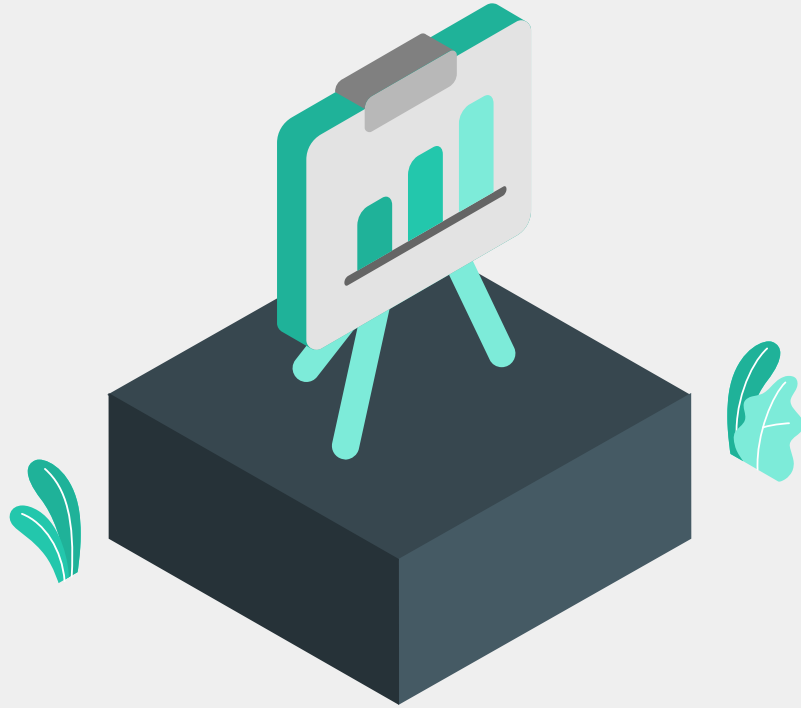
T-TEST ANALYSIS

For hypothesis testing

04

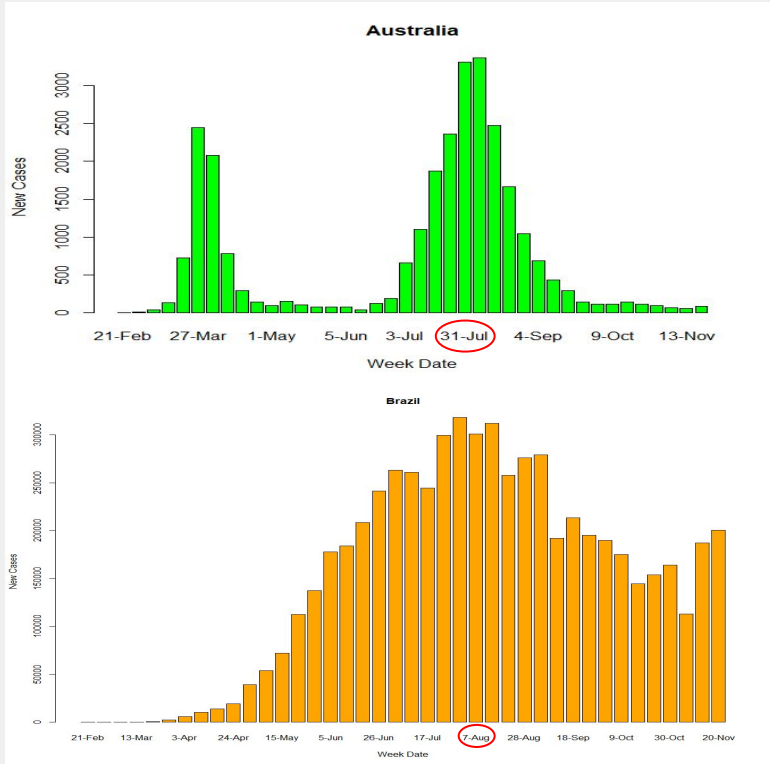
LINEAR REGRESSION ANALYSIS

See how the spotify music features correlate to covid cases.

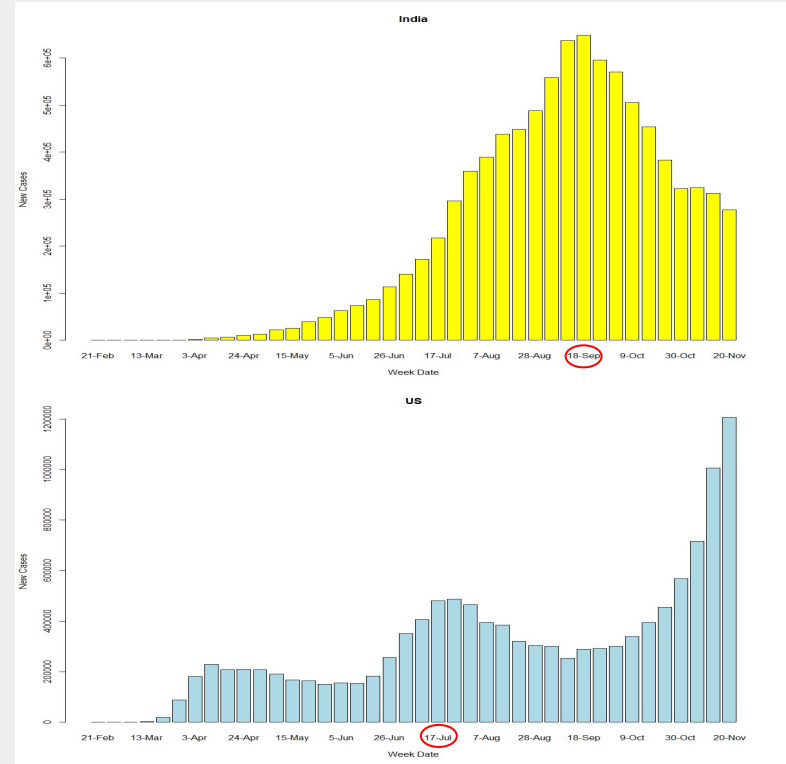


RESULTS

BARPLOTS OF COVID CASES OVER TIME IN 4 COUNTRIES

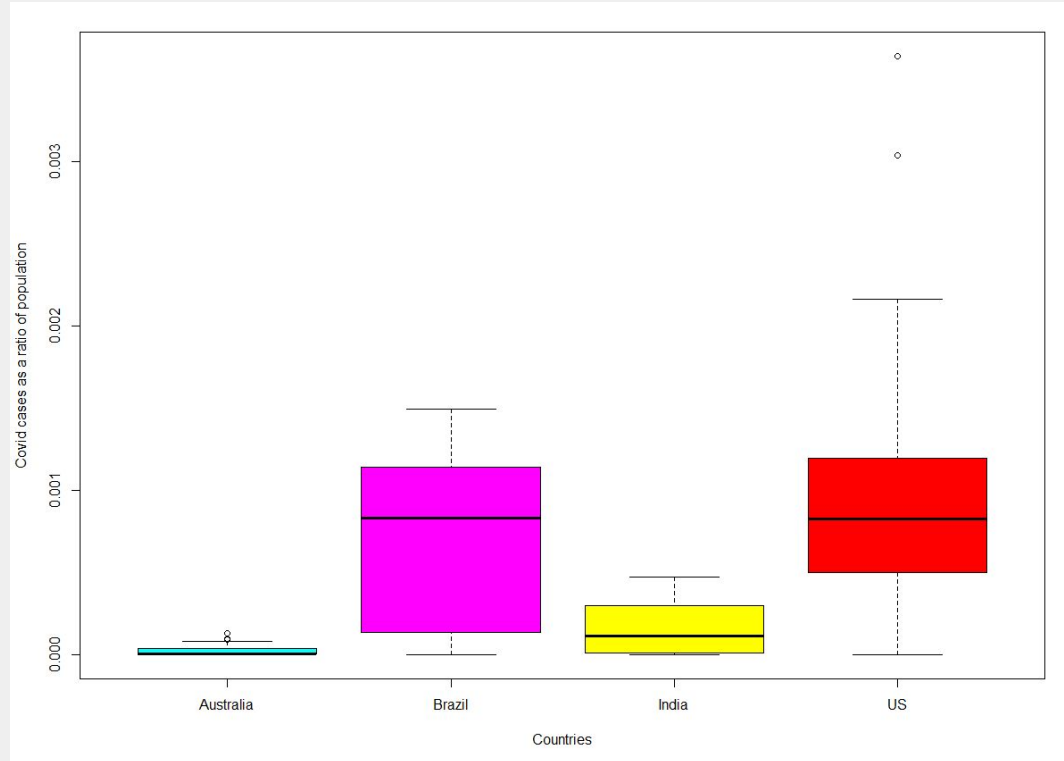


RESULTS - COVID DATA



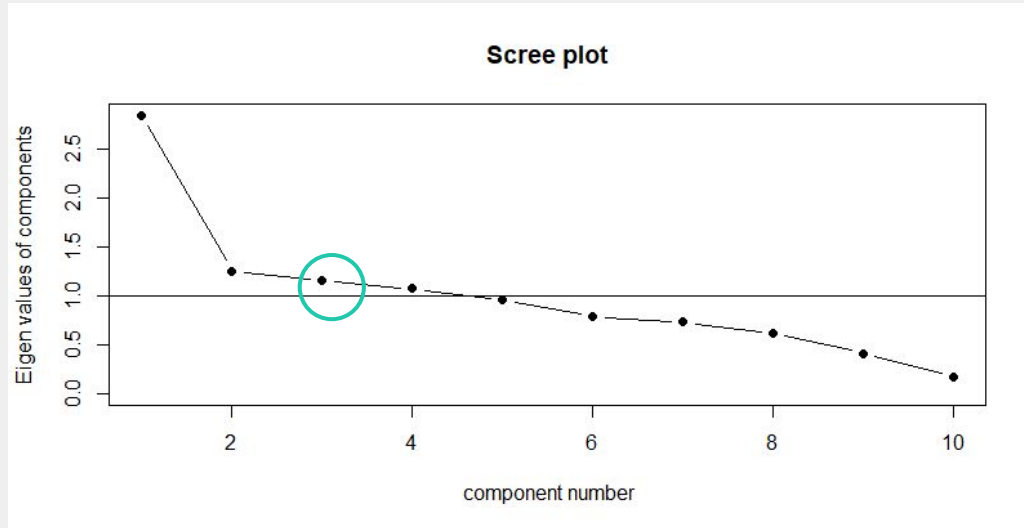
BOXPLOTS OF COVID CASES AS A RATIO OF POPULATION IN EACH COUNTRY

RESULTS - COVID DATA



Australia and India handle COVID relatively well compared to Brazil and USA

RESULTS - PCA



4 centers are chosen for PCA analysis from scree plot

RESULTS - PCA

PCA Analysis	RC1	RC2	RC3	RC4
SS loadings	2.44	1.51	1.30	1.08
Proportion var	0.24	0.15	0.13	0.11
Cumulative Var	0.24	0.40	0.53	0.63
Proportion explained	0.39	0.24	0.21	0.17
Cumulative Proportion	0.39	0.62	0.83	1.00

RC1 which refers to “Energy” and “loudness” explain the most

RESULTS - PCA

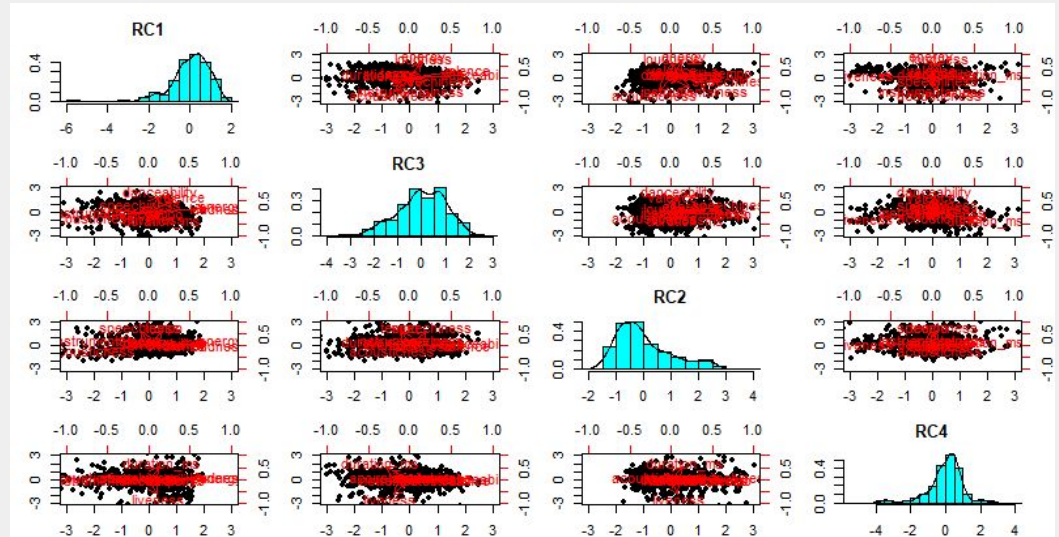
PCA EQUATIONS FOR GENERATING SCORE

$$RC1 = 0.87(\text{energy}) + 0.85(\text{loudness})$$

$$RC3 = 0.83(\text{danceability})$$

$$RC2 = 0.74(\text{speechiness}) + 0.76(\text{tempo})$$

$$RC4 = 0.65(\text{duration_ms})$$

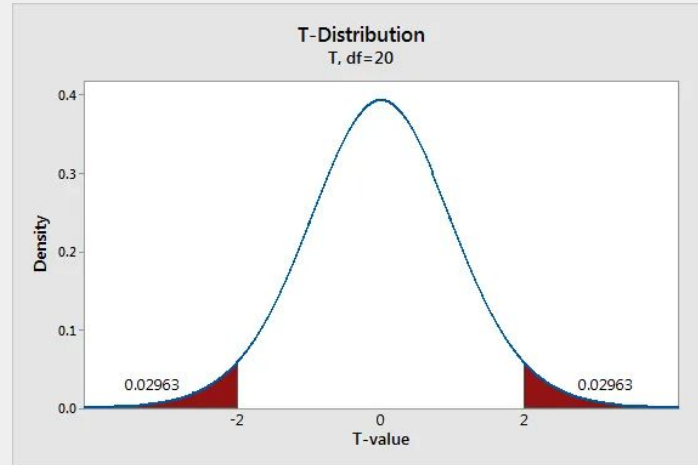


RESULTS - HYPOTHESIS TESTING

Hypothesis testing was done using **two sample t-test**

H₀ : Feature mean didn't change significantly before and after the covid cutting point

H_a : Feature mean was different before the covid cutting point and after the covid cutting point



RESULTS - HYPOTHESIS TESTING

Australia	Before	After
Danceability	0.680	0.676
Energy	0.600	0.622
Valence	0.500	0.498
Liveness	0.174	0.173
Acousticness	0.280	0.267
Speechiness	0.106	0.104

AUSTRALIA

Brazil	Before	After
Danceability	0.675	0.676
Energy	0.724	0.734
Valence	0.648	0.665
Liveness	0.406	0.336
Acousticness	0.343	0.327
Speechiness	0.113	0.116

BRAZIL

India	Before	After
Danceability	0.649	0.647
Energy	0.620	0.620
Valence	0.501	0.495
Liveness	0.167	0.158
Acousticness	0.347	0.348
Speechiness	0.085	0.091

INDIA

US	Before	After
Danceability	0.722	0.695
Energy	0.588	0.608
Valence	0.485	0.464
Liveness	0.182	0.187
Acousticness	0.243	0.232
Speechiness	0.150	0.128

USA

Highlighted rows indicate statistically significant differences between the means

RESULTS - HYPOTHESIS TESTING

Australia	Before	After
Danceability	0.6804926	0.6757338
Energy	0.5958081	0.6215506
Valence	0.5004671	0.4980502
Liveness	0.1737744	0.1729307
Acousticness	0.2797192	0.2667254
Speechiness	0.1061340	0.1041901

India	Before	After
Danceability	0.6485671	0.6470144
Energy	0.6200149	0.6201356
Valence	0.5005559	0.4950676
Liveness	0.1672221	0.1580653
Acousticness	0.3468727	0.3479187
Speechiness	0.08548113	0.09009467

Highlighted rows indicate statistically significant differences between the means

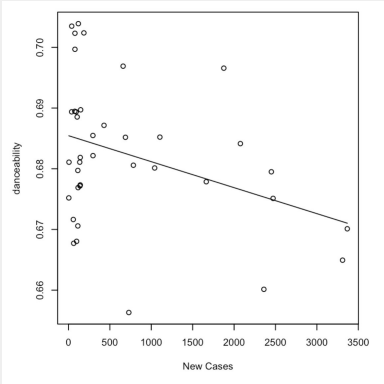
RESULTS - HYPOTHESIS TESTING

Brazil	Before	After
Danceability	0.6750844	0.6761127
Energy	0.7240254	0.7341527
Valence	0.6484552	0.6651967
Liveness	0.4056403	0.3361923
Acousticness	0.3429662	0.3273445
Speechiness	0.1134255	0.1159275

US	Before	After
Danceability	0.7116611	0.6946972
Energy	0.5882287	0.6084817
Valence	0.4851207	0.4642881
Liveness	0.1819240	0.1865472
Acousticness	0.2433358	0.2316420
Speechiness	0.1457597	0.1282322

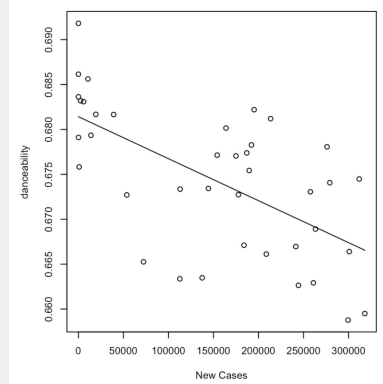
Highlighted rows indicate statistically significant differences between the means

RESULTS - LINEAR REGRESSION



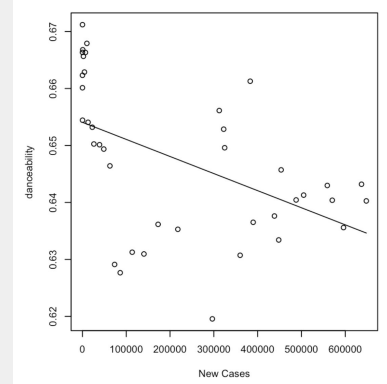
AUSTRALIA

(p-value: 0.02702)



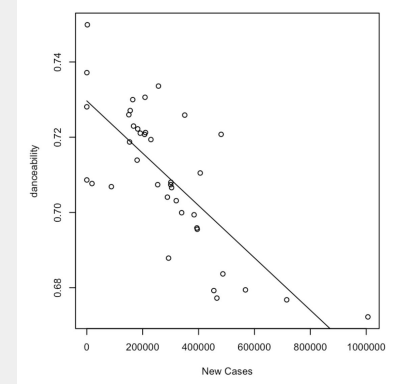
BRAZIL

(p-value: 0.000022)



INDIA

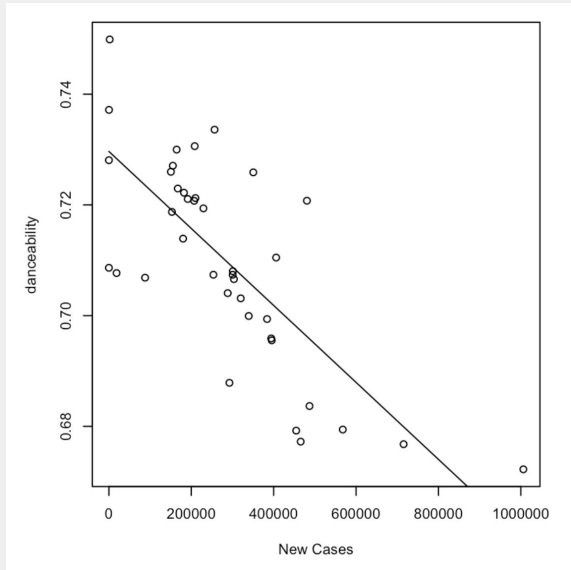
(p-value: 0.001233)



USA

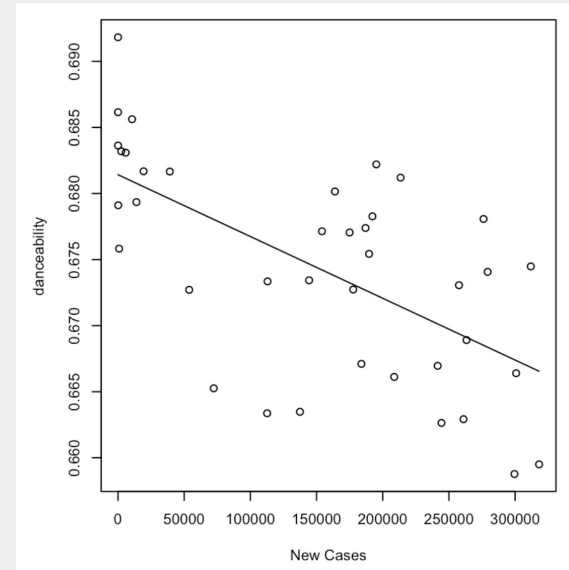
(p-value: 0.000000033)

RESULTS - LINEAR REGRESSION



USA

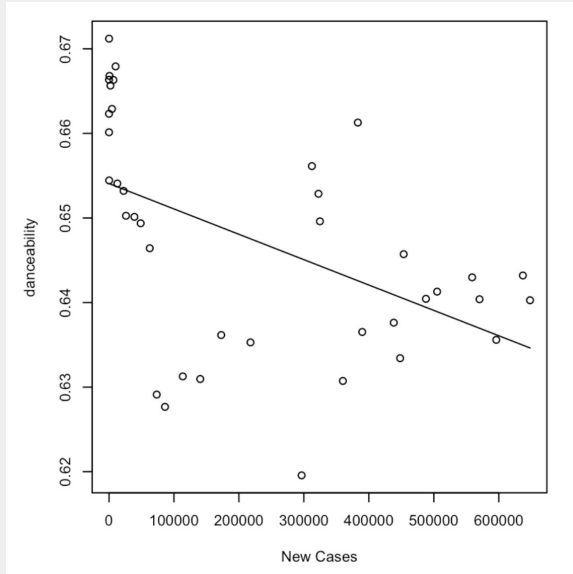
(p-value: 0.000000033)



BRAZIL

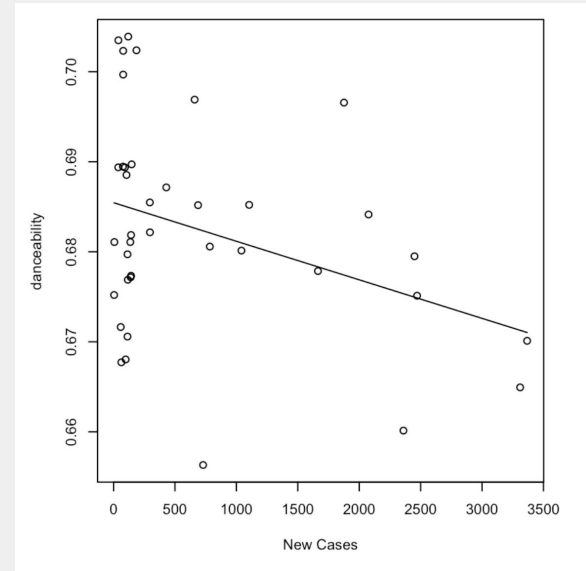
(p-value: 0.000022)

RESULTS - LINEAR REGRESSION



INDIA

(p-value: 0.001233)



AUSTRALIA

(p-value: 0.02702)



LIMITATIONS & ACKNOWLEDGEMENT

LIMITATIONS

DATA

- How Spotify gets the feature values is unknown
- Limited access on data dimension

OTHER FACTORS

- Change could be affected by other factors which we have not accounted for
- COVID cutoff date might not be entirely accurate

ACKNOWLEDGEMENT

“I hope you consider the statistical test to examine the significant difference about music trend..”

- Added hypothesis testing

“Maybe segment the data to multiple subsets..”

- Attempt, but not effective due to the sheer diversity existent
- We segmented by countries to try to see correlations between case numbers and music features

SUMMARY

01

COVID HAS CHANGED MUSIC TRENDS

Data dimension was reduced though PCA , verified by t-test

02

THE WORSE A COUNTRY HANDLES COVID, THE MORE FEATURES CHANGE

analyzed with PCA and hypothesis testing

03

THE BETTER A COUNTRY HANDLES COVID, THE LESS CORRELATED ARE MUSIC FEATURES AND COVID NUMBERS

analyzed with Linear Regression

THANK YOU!

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik** and illustrations by **Stories**

Please keep this slide for attribution.