# Spotify Streams

Team:
Jorge Chavez
Jessica McPhaul
Dominic Marin

# Essential Mission:

We are assisting the music industry to create top streaming music. These findings will help determine any trends that help produce a top streaming music/songs.

# Research questions:

1. What features have been the most prevalent in the past 10 years for a top streaming music/song?
2. What is the association of a music artist's number of streams with their fame online?
3. What is the correlation between disability and the popularity of top 5 songs in the last 10 years? (To determine if the music is more danceable makes it more widespread.)

# Data Exploration and Cleanup process

1.Finding relevant data

2.Organize data so it could be imported, read and analyzed

3.Formatted column to merge

4.Convert data to variables

5.Create 5 different data frames from which to analyze data

6.Eliminate void/null/unnecessary columns

DATASOURCES:

KAGGLE

# Question 1: What features have been the most prevalent in the past 10 years for a top streaming music and songs?
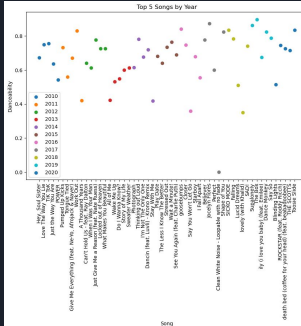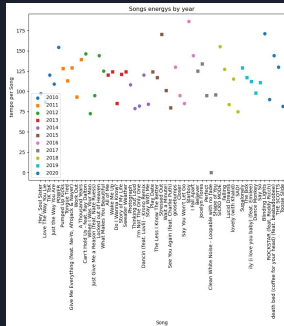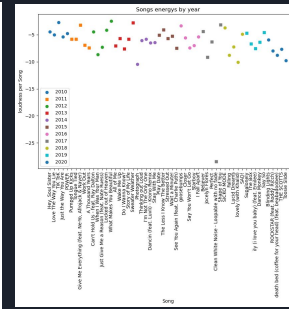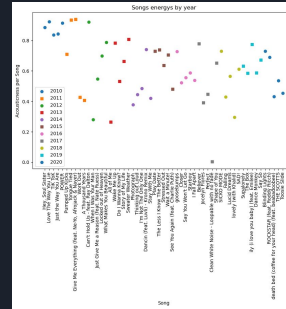
Features used to gauge top streaming songs per year:

- acousticness, danceability, energy,liveness, loudness, and tempo : was not able to find Not able to find strong correlations between popularity and song features.

Results

- This maybe because a song popularity will be surrounded with what's going on during that year such trends.
- Also each Top 5 song Per Year, were all in different genres, so it will affect the audio features of each song uses

# Audio Features Visualizations



In Conclusion, audio features were not a conclusive way to determine what made a song hit within each year because of weak correlation.
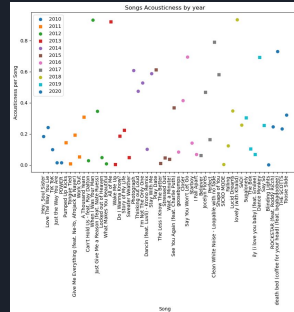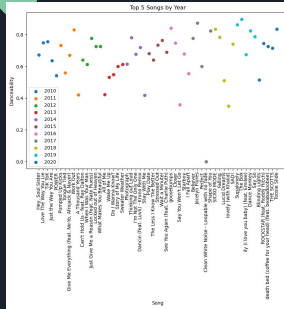
|  | acousticness | danceability | energy | liveness | loudness | popularity | tempo |
|---|---|---|---|---|---|---|---|
| count | 55.000000 | 55.000000 | 55.000000 | 55.000000 | 55.000000 | 55.000000 | 55.000000 |
| mean | 0.281603 | 0.664382 | 0.614680 | 0.187515 | -6.462691 | 85.818182 | 113.106036 |
| std | 0.270509 | 0.160450 | 0.196764 | 0.164359 | 3.566008 | 5.891271 | 30.334678 |
| min | 0.001460 | 0.000000 | 0.003420 | 0.059600 | -28.460000 | 78.000000 | 0.000000 |
| 25% | 0.057650 | 0.599500 | 0.467500 | 0.093700 | -7.421500 | 82.000000 | 94.003000 |
| 50% | 0.192000 | 0.689000 | 0.631000 | 0.123000 | -5.849000 | 84.000000 | 116.879000 |
| 75% | 0.441500 | 0.770000 | 0.740500 | 0.206500 | -4.717500 | 90.000000 | 129.000000 |
| max | 0.934000 | 0.896000 | 0.939000 | 0.790000 | -2.494000 | 100.000000 | 186.003000 |

# Question 2: What is the association of a music artist's number of streams with their fame online?

Of the top streamed songs, 60% of them were from artists that have a huge following online. Instagram has the largest correlation between followers and spotify streams, with Youtube coming in behind it.

Over the last 3 years, artists whose songs hit number one on Spotify saw at least a 32% increase in followers and subscribers in their respective Instagram accounts and Youtube channels.

And artists who have amassed a huge online presence within the two outlets, have had more than double the streams of artists who do not.

There is a huge incentive to prioritize social media and gain a following, as it has a positive correlation the number of streams on Spotify.

One major difference is that, artists can monetize their media platforms while the royalties from Spotify streams are miniscule.  They generate income from social media, while they generate followers from their Spotify.

# The Process

| | | | | |
|---|---|---|---|---|
| data | ✓ | 2/15/2023 4:25 PM | File folder | |
| Project_one | ✓ | 2/15/2023 5:15 PM | File folder | |
| project1 | ⟳ | 2/14/2023 7:31 PM | File folder | |
| project-one | ⟳ | 2/15/2023 2:18 PM | File folder | |
| working_dupes | ✓ | 2/15/2023 1:00 AM | File folder | |
| .vscode-janus-debug | ✓ | 2/15/2023 12:31 AM | VSCODE-JANUS-DEBUG... | 0 KB |
| completenotebook.ipynb | ✓ | 2/15/2023 5:32 PM | Jupyter Source File | 7,515 KB |
| kaggle.json | ✓ | 2/13/2023 5:04 PM | JSON Source File | 1 KB |
| kaggle_api.json | ✓ | 2/13/2023 5:06 PM | JSON Source File | 1 KB |
| project_one.ipynb | ✓ | 2/15/2023 2:12 AM | Jupyter Source File | 11 KB |
| project1.ipynb | ✓ | 2/15/2023 12:41 AM | Jupyter Source File | 955 KB |
| project1.pk1 | ✓ | 2/15/2023 1:56 AM | PK1 File | 398 KB |
| project1.pk2 | ✓ | 2/14/2023 11:18 PM | PK2 File | 800 KB |
| project1.pk3 | ✓ | 2/14/2023 11:11 PM | PK3 File | 3 KB |
| project2.pk2 | ✓ | 2/15/2023 2:07 AM | PK2 File | 7 KB |
| project3.pk3 | ✓ | 2/15/2023 1:03 AM | PK3 File | 7 KB |
| project4.pk4 | ✓ | 2/15/2023 1:04 AM | PK4 File | 2 KB |
| project5.pk5 | ✓ | 2/15/2023 1:04 AM | PK5 File | 15 KB |
| spotify.ipynb | ✓ | 2/14/2023 10:29 PM | Jupyter Source File | 233 KB |
| tmi.pk1 | ✓ | 2/14/2023 10:09 PM | PK1 File | 2 KB |
| topinfluencers.ipynb | ✓ | 2/15/2023 2:07 AM | Jupyter Source File | 353 KB |
| youtube_artists.csv | ✓ | 2/15/2023 4:47 PM | CSV File | 12 KB |

| Name | Status | Date modified | Type | Size |
|---|---|---|---|---|
| online_presence.csv | ✓ | 2/15/2023 12:34 AM | CSV File | 1 KB |
| project_one.ipynb | ✓ | 2/15/2023 2:14 AM | Jupyter Source File | 1 KB |
| spotify_dataset.csv | ✓ | 2/14/2023 10:28 PM | CSV File | 437 KB |
| Spotify_final_dataset.csv | ✓ | 2/15/2023 2:04 AM | CSV File | 644 KB |
| spotify_streams.csv | ✓ | 2/14/2023 11:22 PM | CSV File | 5 KB |
| spotify1.csv | ✓ | 2/15/2023 1:59 AM | CSV File | 437 KB |
| streamed_the_most.csv | ✓ | 2/14/2023 11:21 PM | CSV File | 5 KB |
| tmi.csv | ✓ | 2/14/2023 10:08 PM | CSV File | 1 KB |
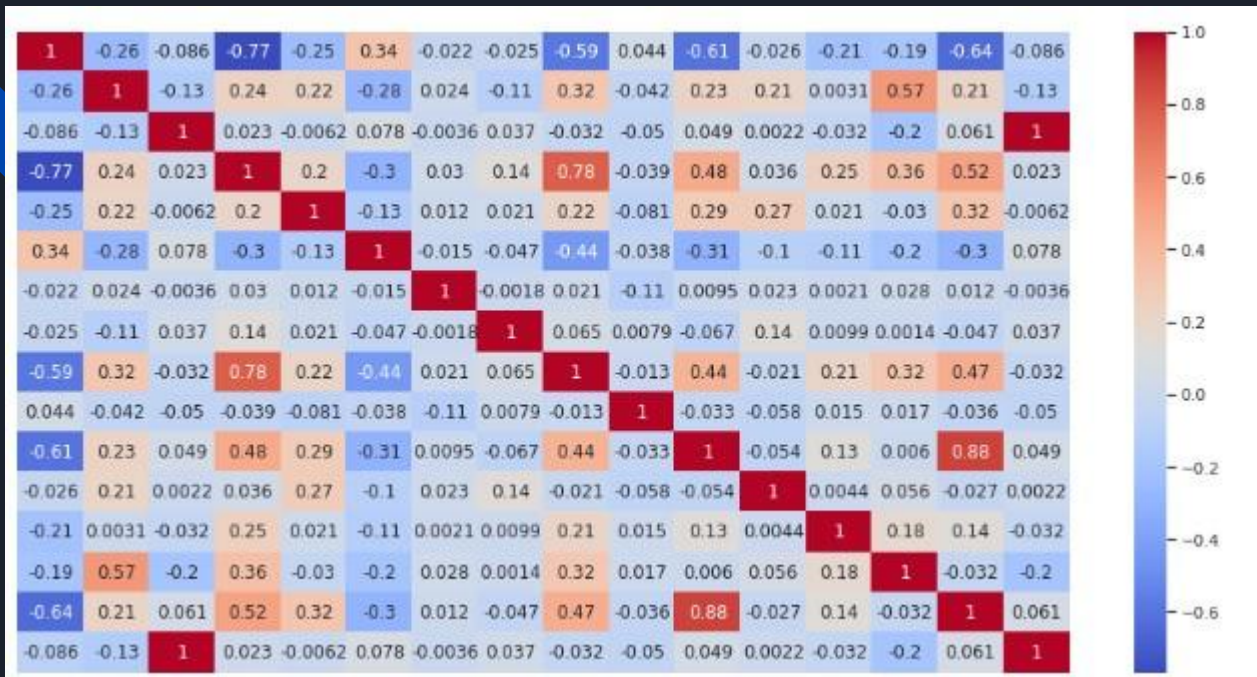
Correlation analysis between combined Youtube and Instagram followers per artist and their spotify streams and expected streams based on the number of their followers

# Precision Recall From Complete Data.  Close to 1.

- *Precision-Recall Score:* **0.9096**

- *Precision-Recall Curve:*
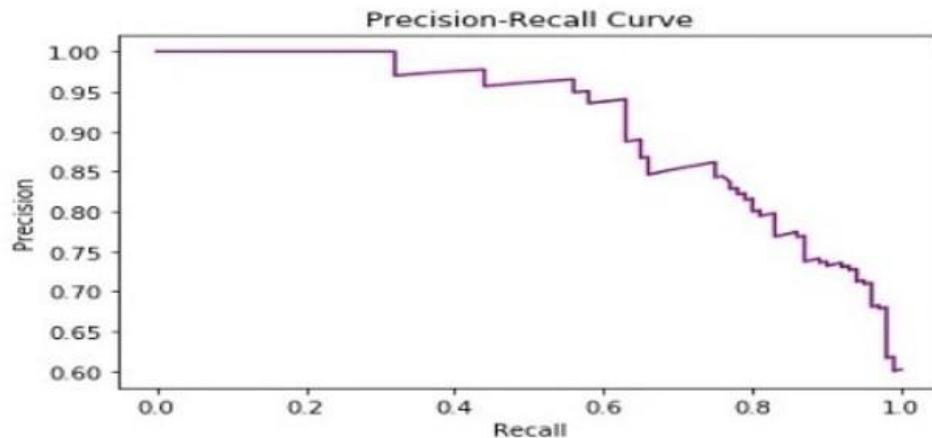
Question 3: What is the correlation between the danceability and the popularity of top 5 songs in the last 10 years? (To determine in the music is more danceable makes it more widespread)

# Additional Resources

https://www.thefamouspeople.com/musicians.php

https://towardsdatascience.com/top-music-analysis

https://www.kaggle.com/dataset/starter-music-artists-popularity

https://soundcharts.com/blog/music-analytics-tools

https://www.kaggle.com/code/varunsaikanuri/spotify-data-visualization/data

https://datascientyst.com/science-as-music/business

https://www.visualcapitalist.com/worlds-top-50-influencers-across-social-media-platforms/

# Conclusion

In our presentation, were we able to see:

- As an artist or institution that may use or have their work on a music streaming platform, it is important to understand the vast music streaming landscape and options
- to find the service that best fits one's individual needs.
- For music streaming companies, it is essential to know what competitors are offering and new trends in this ever-changing model.
- In the second part of this research, we will present recommendations for a business seeking to stay competitive in this evolving industry.
- For artists, increasing their online presence means increasing streams and vice-versa.

# Questions?