2010-2022 Spotify Top Hits Analysis

Cody Allison

Hannah Weber

James Beckmeyer

Summary of Proposal

For our proposal we analyzed the top hits from Spotify spanning from the 2010s through 2022. Using Spotify's API data we extracted track attributes or traits for each top hit. We then identified if those traits have changed over the decade, and if the Covid outbreak in 2020 caused a shift in trends.

Our primary datasets for 2010-2019 came from Kaggle, which pulls from official Spotify playlists for top hits of those years. These playlists were made of 100 total tracks each. For 2020 and 2022, we will be pulling directly from similar Spotify playlists for those years and using the Spotify API to extract the Track attributes. These playlists, unlike the previous years, only included 50 tracks total. 2021 did not have an official playlist, so we will be using a separate Kaggle dataset that compiled the 50 most streamed songs of that year.

Metrics

To examine trends, we used metrics defined and measured by Spotify. A summary of each attribute can be found within Spotify's API documentation. These definitions have been copied below:

Acousticness: A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.

Danceability: Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.

Energy: Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.

Valence: A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

Speechiness: Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.

Popularity: The popularity of a track is a value between 0 and 100, with 100 being the most popular. The popularity is calculated by algorithm and is based, in the most part, on the total number of plays the track has had and how recent those plays are.
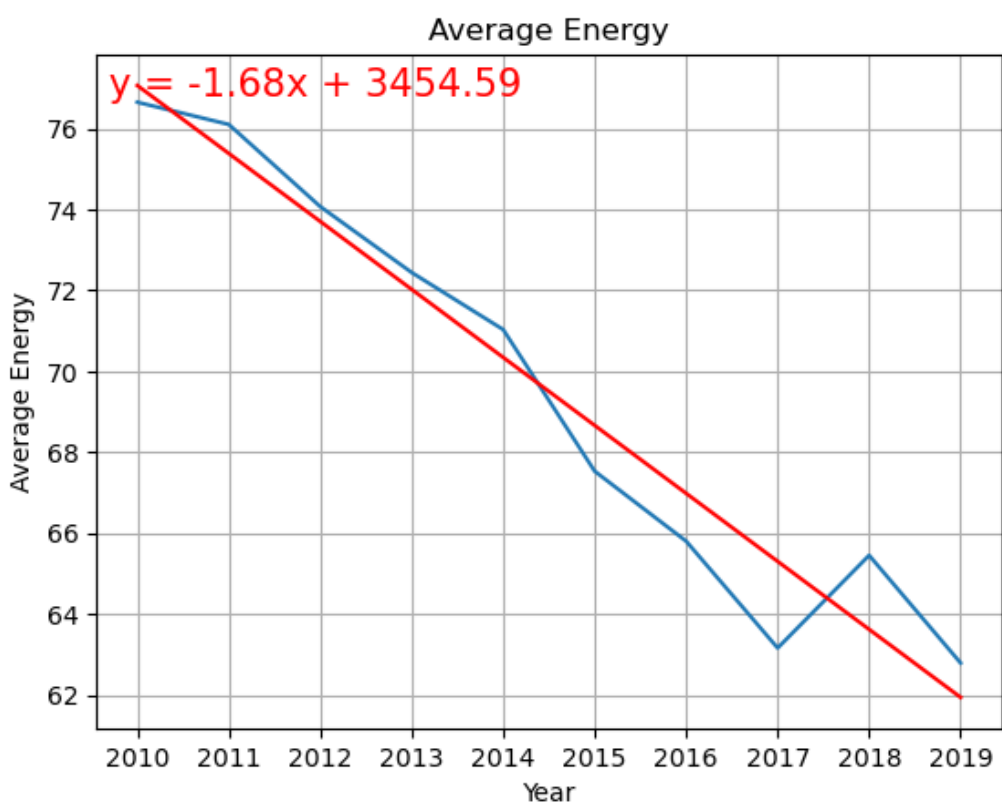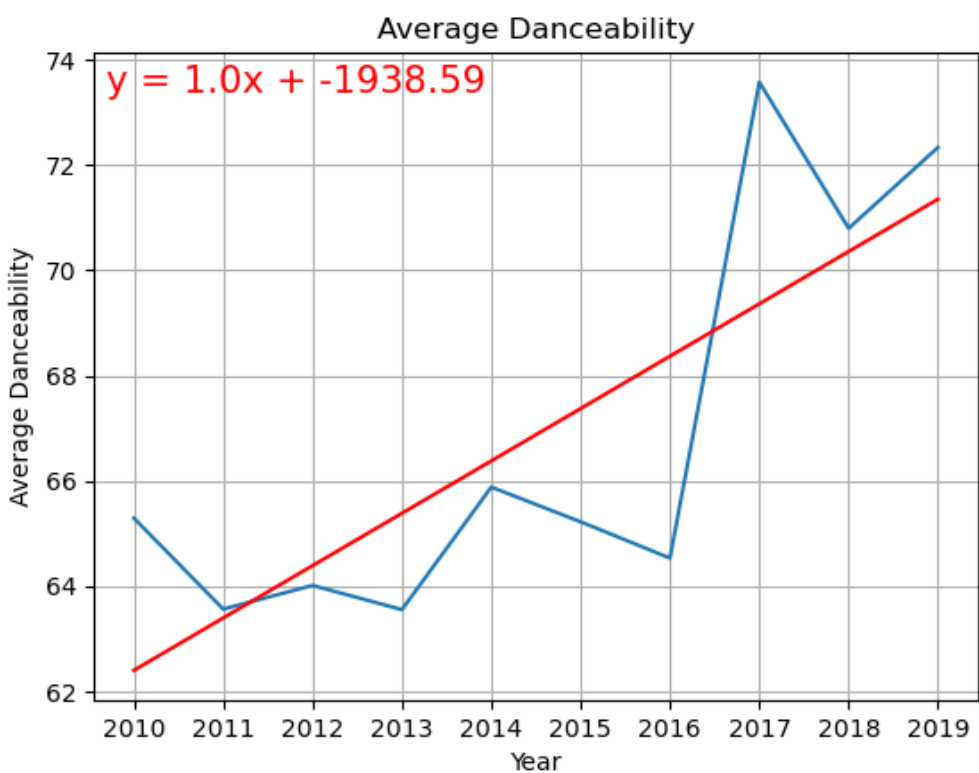 Generally speaking, songs that are being played a lot now will have a higher popularity than songs that were played a lot in the past. Duplicate tracks (e.g. the same track from a single and an album) are rated independently. Artist and album popularity is derived mathematically from track popularity. ***Note***: *the popularity value may lag actual popularity by a few days: the value is not updated in real time.*
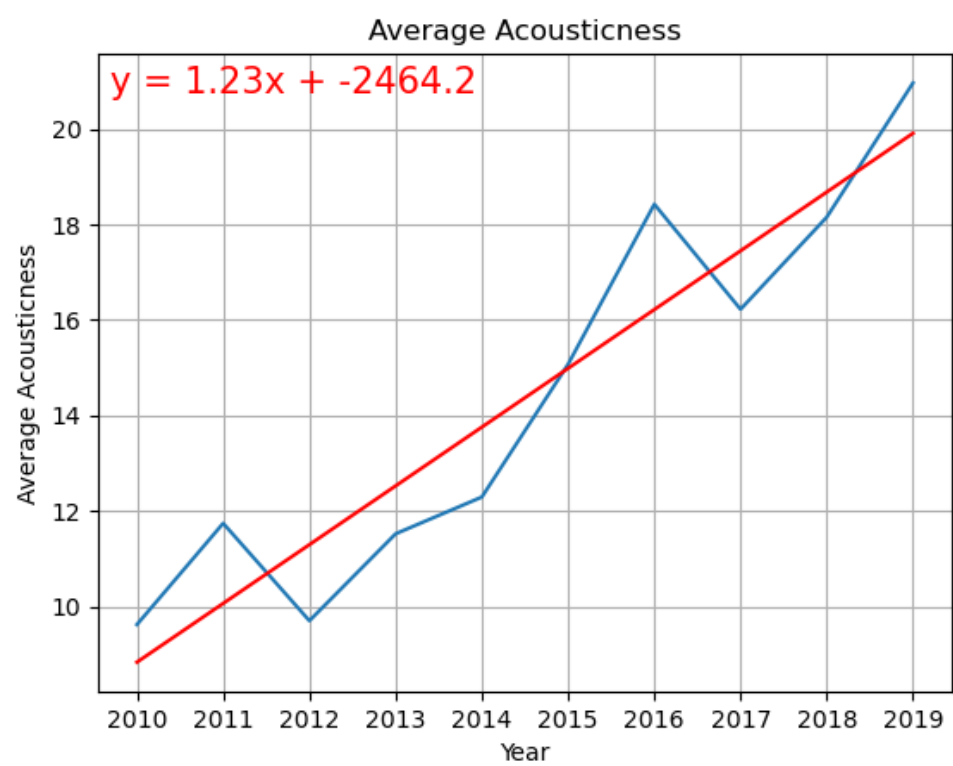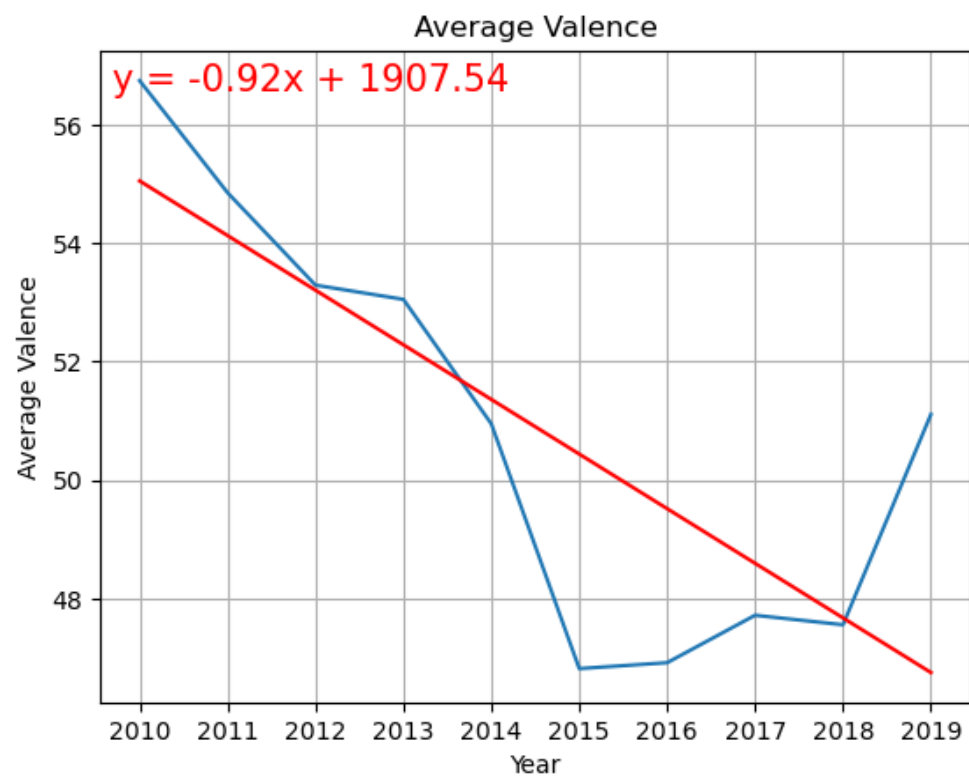
What traits are present in top hits from the 2010s and did they change over the decade?
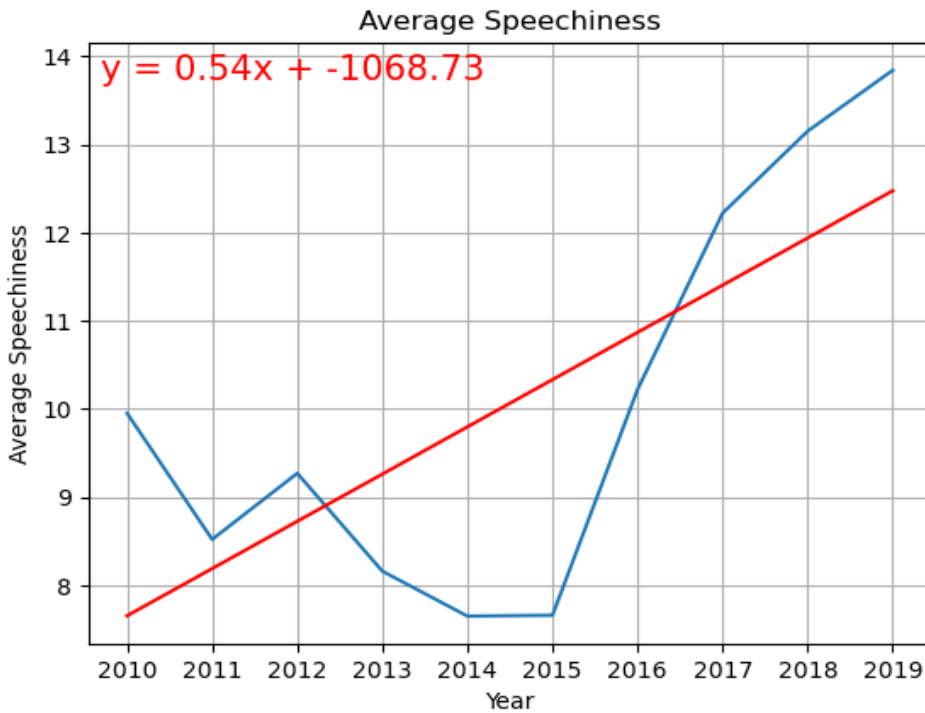

Average Track Attribute over 2010s
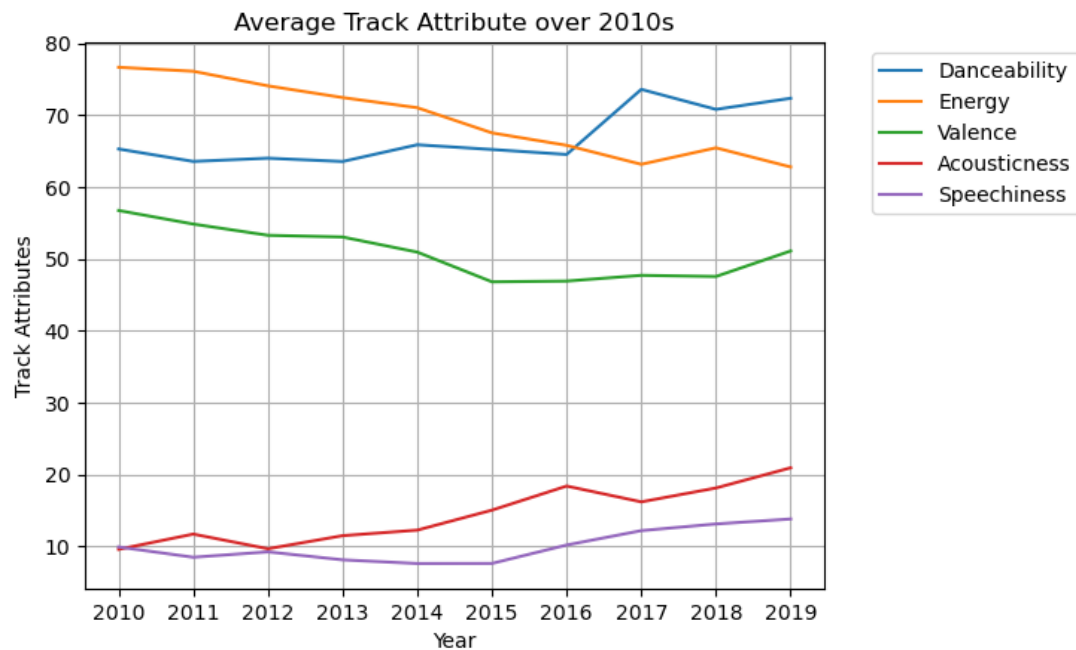
Above we have taken the average of our 5 primary attributes and plotted them throughout the decade. Acousticness, speechiness and danceability generally seem to rise. Energy appears to fall over time, and valence ultimately appears to trend slightly lower then level off. To better visualize each metric individually, we created linear regressions to display the trends (see below). To test our hypothesis that "the average track attribute follows a trend through the 2010s", we established a null hypothesis that "each trait ultimately does not follow a trend throughout the 2010s". We then tested this hypothesis against our 5 primary trends using a-nova tests, giving a p-value for each attribute. Using the data below, we can reject the null hypothesis for all traits aside from energy. This indicates each trait is potentially following the trend presented in the graphs below.
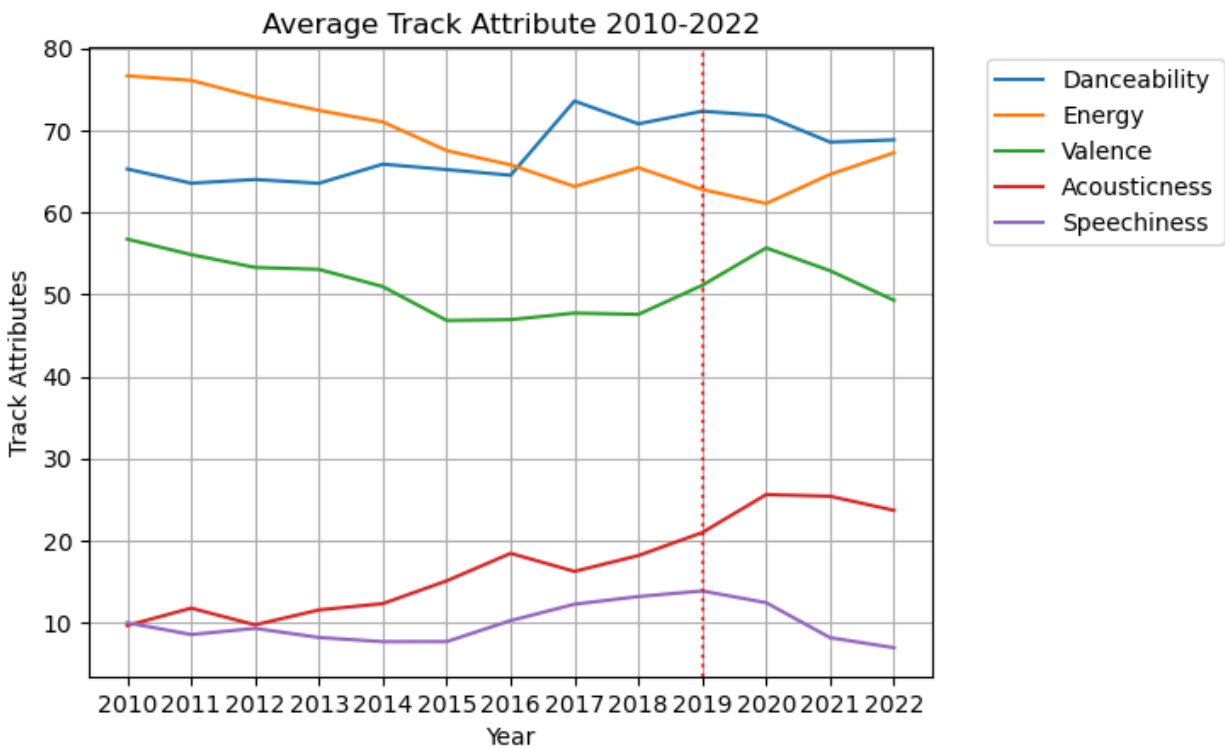
2010s P-Values	
Trait	P-Value
Danceability	2.06e-5
Energy	.19
Valence	.1.05e-17
Acousticness	1.77e-05
Speechiness	9.76e-09

Average Danceability

y = 1.0x + -1938.59

Average Energy

y = -1.68x + 3454.59

Average Valence

y = -0.92x + 1907.54

Average Acousticness

y = 1.23x + -2464.2
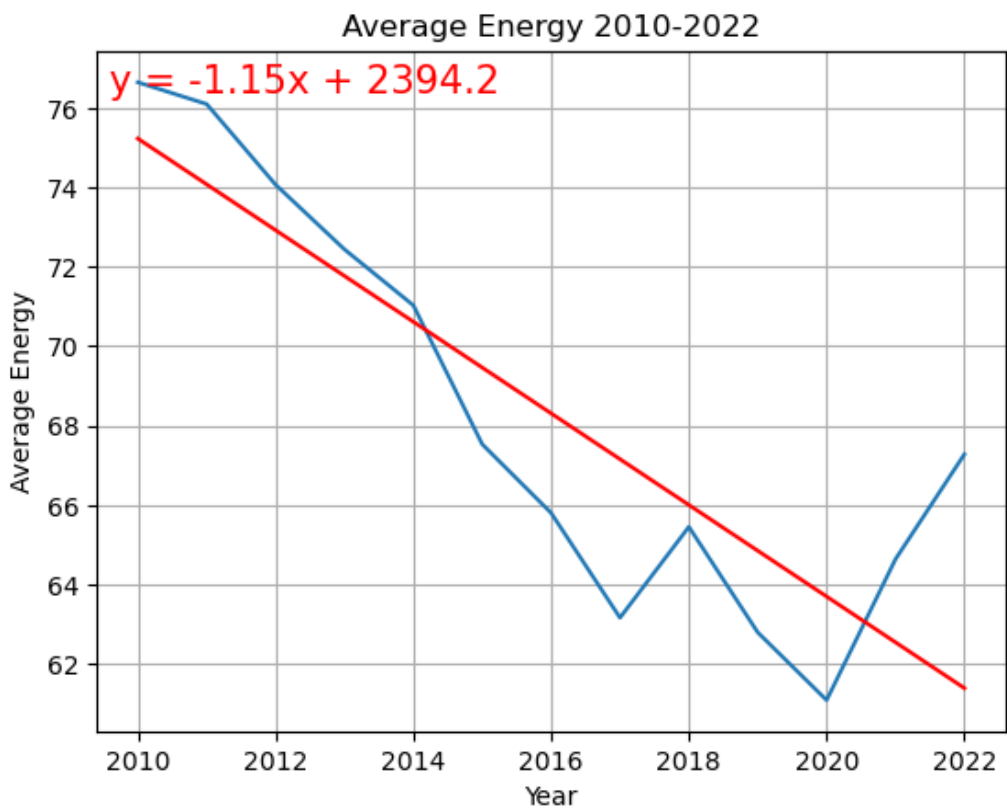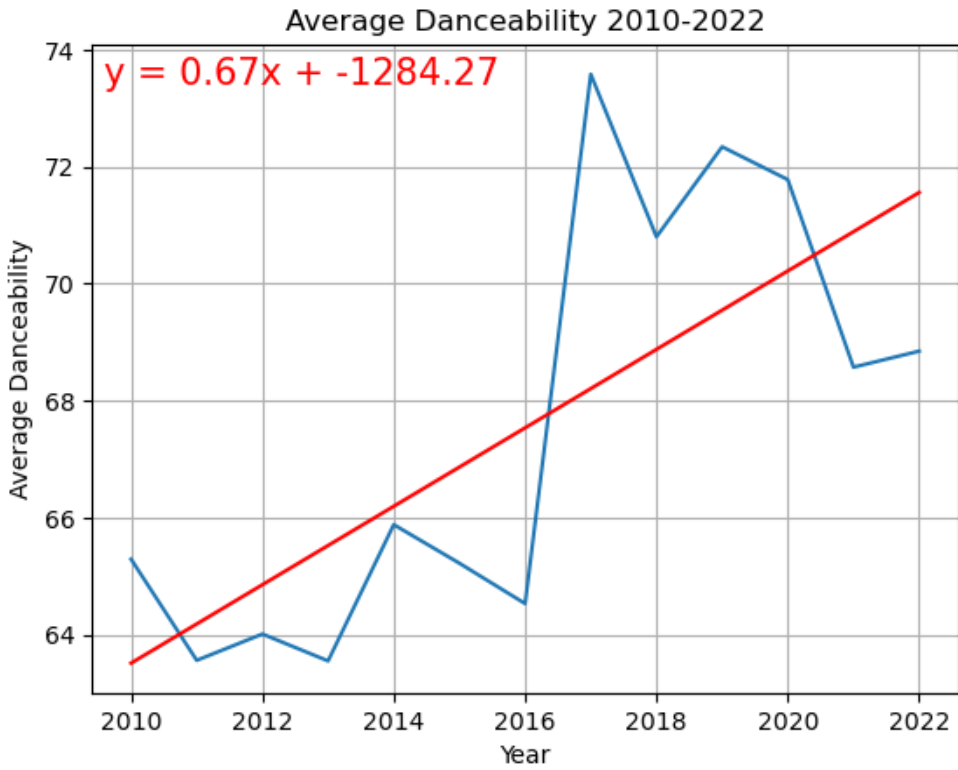
Average Speechiness

$y = 0.54x + -1068.73$

Was the Covid outbreak a cause for a large shift in trends?

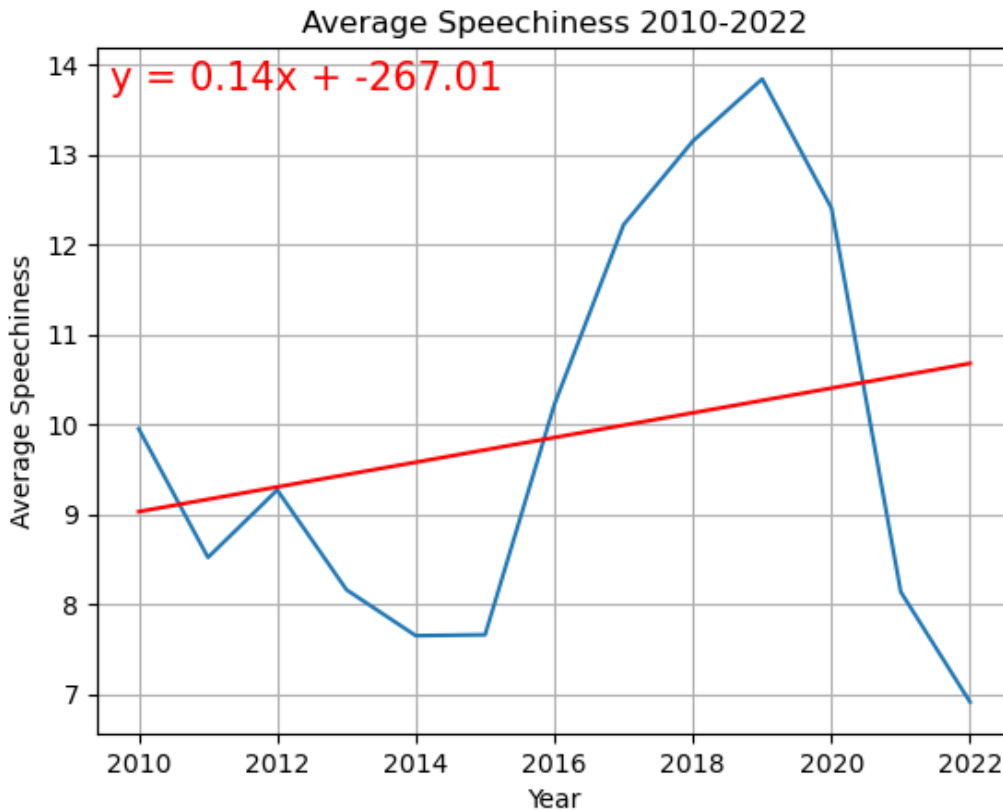Average Track Attribute 2010-2022

Above we have taken the average of our 5 primary attributes and plotted them from 2010 to 2022. Average acousticness, speechiness and danceability seem to rise. Energy appears to fall, and valence ultimately appears to dip then level off. To better visualize each metric individually, we created linear regressions to display the trends (see below). To answer our question of "was the Covid outbreak the cause for a large shift in trends, we outlined an alternative and null hypothesis. The alternative hypothesis was that the average track attributes do follow a trend from the 2010s to the 2020s. The null hypothesis was that the average track attributes do not follow a trend from the 2010s to 2020s. To see if the null hypothesis was applicable, a t-test was performed to compare the 2010s trait values to the 2020s trait values. If the null hypothesis is not applicable, this means there is a possibility there is a trend from the 2010s to the 2020s, so there is not evidence that the Covid outbreak caused a large shift in trends. The traits examined were danceability, energy, valence, acousticness, and speechiness. For danceability, energy, and acousticness, the null hypothesis can be rejected due to the size of the p-value being less than the threshold of 5%. From the available information, there is no evidence to indicate the average track attributes do not follow a trend from the 2010s to 2020s, so the Covid outbreak did not cause a shift in trends for these traits. For valence and speechiness, the null hypothesis cannot be rejected due to the size of the p-value being greater than the threshold of 5%.. This means there is a possibility the average track attributes do not follow a trend from the 2010s to the 2020s, so the Covid outbreak could have caused a shift in trends. Further analysis would still be needed to prove the Covid outbreak was the cause of the shift in trends.

2010s vs 2020s T-Test	
Trait	P-Value
Danceability	.0121
Energy	7.68e-5
Valence	.3967
Acousticness	1.58e-6
Speechiness	.2218

Average Danceability 2010-2022

y = 0.67x + -1284.27

Average Energy 2010-2022

y = -1.15x + 2394.2

Average Valence 2010-2022

y = -0.31x + 679.87

Average Acousticness 2010-2022

y = 1.41x + -2832.48

Average Speechiness 2010-2022

y = 0.14x + -267.01
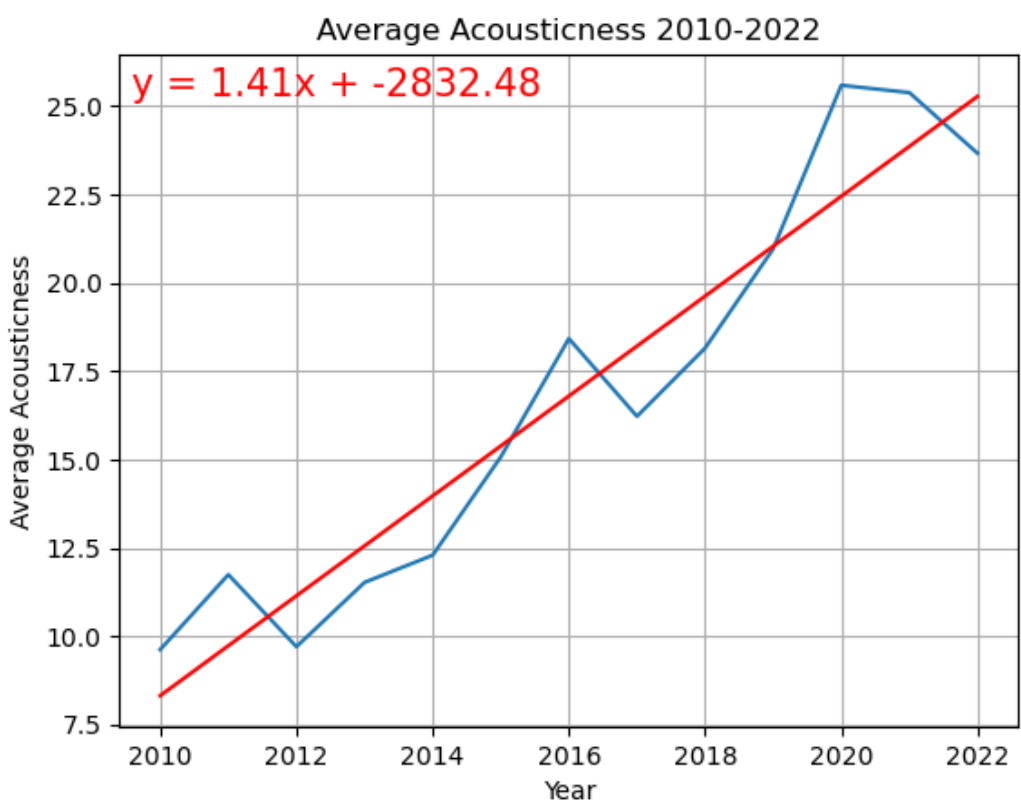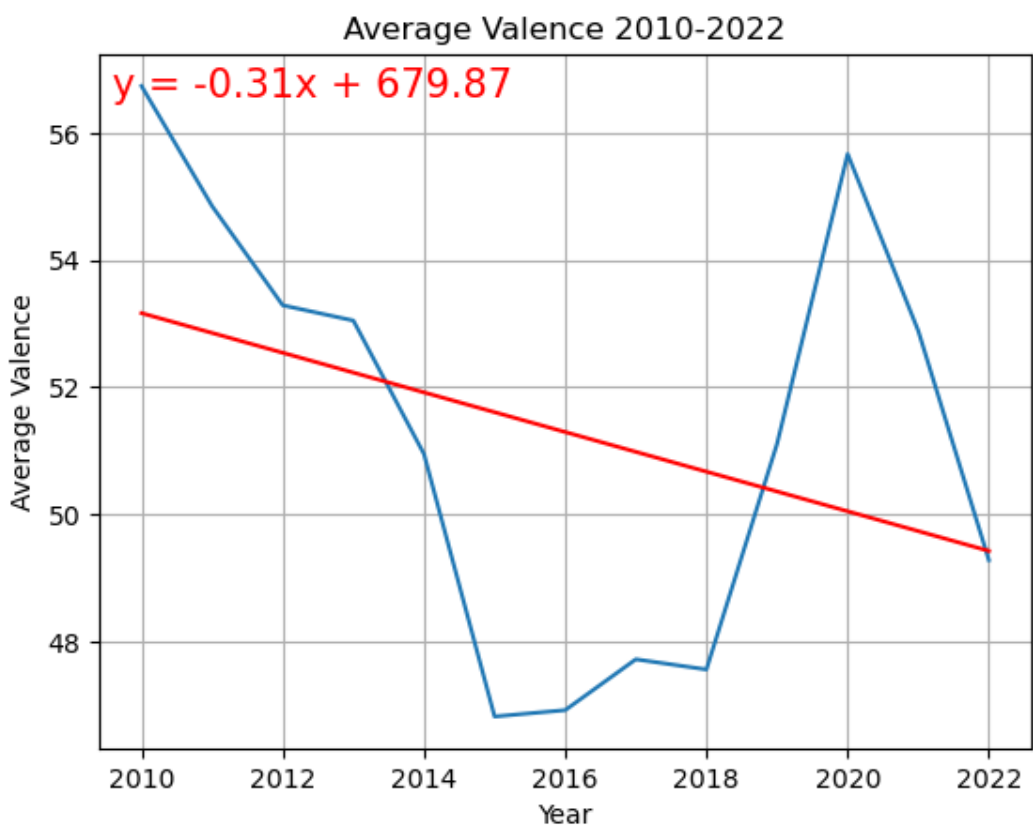
For more information on our T-Tests, ANOVA, statistical analysis, please refer to the jupyter notebook file **all_years_playlist_data.ipynb**.

Data Sources:

https://www.kaggle.com/datasets/muhmores/spotify-top-100-songs-of-20152019

https://www.kaggle.com/datasets/equinxx/spotify-top-50-songs-in-2021

Directly pulling from Spotify

Spotify API documentation:

https://developer.spotify.com/documentation/web-api/reference/get-several-audio-features

Accessing the Spotify API with Python
https://lvngd.com/blog/accessing-spotify-api-python/