

We Are All Trying to Describe Something And We Don't Have the Words*

A Study of The Jack Antonoffication of Pop Music (Or: Propagating Uncertainty with Spotify)

Michaela Drouillard

4/19/23

Using data from the Spotify API, I explored Jack Antonoff's production style and assessed whether Spotify's audio features can capture it. In an interview with Glen McDonald, an engineer at Spotify, I learned that the danceability and valence features were created by having interns tag songs as positive or gloomy, or danceable or un-danceable. Using a logistic regression model, I found that danceability was the most important feature in predicting Antonoff's involvement in a song. In the discussion section, I do a close-listen of Antonoff's most danceable songs, and argue that the danceability variable is actually capturing something more like a visceral quality to Antonoff's music.

1 Introduction

Sometime in early October of 2022 there was a chemical shift in my brain and I couldn't stop listening to Taylor Swift. I couldn't listen to anything other than Fearless or Red when deadlifting, walking to a friend's house, folding laundry, etc. I listened to so much Taylor in that month alone that she ended up as my #4 artist on Spotify Wrapped, which stops collecting its data for the year on October 31st. It didn't even capture what happened to me after *Midnights* dropped in November.

I was standing outside of the Garrison eating a slice of pizza with some friends, and one of them asked what music we had been listening to lately.

ME: "Midnights!"

ZACH: "Is it good?"

*Code and data are available at: <https://github.com/michaeladrouillard/jack>.

[I glitched.]

ME: “No!”

[“No?”]

A STRANGER WALKING BY WHO HEARD EVERYTHING: “EXACTLY!!”

Midnights was produced by Swift and her longtime collaborator Jack Antonoff, who is the subject of this paper, and, I suspect, the reason behind both my glitching, and the stranger’s “Exactly”.

Antonoff is prolific, to the extent that it is starting to bother people. He has produced and co-produced music for Lorde, Taylor Swift, Florence and the Machine, The Chicks, Clairo, The 1975, Grimes, Zayn, Pink, Lana Del Rey, Olivia Rodrigo, the Minions: The Rise of Gru Soundtrack, and more. All of this and his own bands too (Bleachers, and fun.). When I was getting brunch with a friend one Sunday and a Swift song came on, and she said “Jack Antonoff’s gonna kill us all”, we both laughed. What my friend was referring to is the fact that Jack Antonoff and his sound seems to be everywhere these days. As Andrew Marantz writes in the New Yorker: “When his band releases an album, the world responds politely. When he produces one by Lorde or Lana Del Rey or Taylor Swift, the world wobbles on its axis” (Marantz 2022).

What I am trying to get to here is that I don’t actually really mind Jack Antonoff’s music, I kind of like it, but I still say his voice with the same tone that everyone else does. There’s just something about him... this sense that he’s everywhere... and that he’s just such a nice guy... Why do music people seem to be a little annoyed with him? What is it about Antonoff music that is so Antonoff-y?

To find an answer to this question, I turned to statistics, people who know more than me about describing music (Youtubers), and the overlord of engineered virality itself – the Spotify API. Can Spotify’s audio features dataset help us determine what makes Antonoff music so Antonoff-y?

I spoke to Caleb Gamman, a Youtuber and Antonoff-production-style hater who had a tweet go viral around the time of Taylor Swift’s *Midnights* release, in which he predicts whether or not a song was produced by Antonoff within seconds. A real-life logistic regression, if you will. Using an actual logistic regression model, I was able to determine that Spotify’s **danceability** feature was the most important in predicting whether or not a song had any Antonoff involvement. During an interview with Glen McDonald, an engineer at Spotify, I learned that the **valence** and **danceability** variables were created by hiring college interns and having them tag whether or not songs were positive or danceable. In the discussion section, I listen to the songs with the highest danceability scores, and make an argument that this **danceability** feature, which has so much uncertainty cooked into it, is really capturing something like visceral-ness. Against the odds, we can capture something about Antonoff’s music with these variables, and it might just be that subjective “ick” that some people feel, and others enjoy, and that we’re all just trying to describe.

2 An Interview with Caleb Gamman

Around the time that *Midnights* came out, I came across a video on Twitter of a guy, Caleb Gamman, guessing which songs on *Midnights* were produced by Antonoff within seconds (@calebgamman 2022). The tweet reads: “im able to instantly detect if jack antonoff worked on a song due to a visceral hatred of his production style”. He nails it. Only one song on the album wasn’t produced by Antonoff, and he guesses it.

I reached out to Gamman to ask him to explain Antonoff to me, only to find out, on the morning of the interview, that Gamman had released a 29 minute video essay on his Youtube Channel covering (Gamman 2022):

(you can skim the quote chunks)

1. what makes an Antonoff song so Antonoff-y

...I’m literally not an expert, and of course I don’t actually claim to always be able to be able to 100% identify a Jack Antonoff song. It’s not like I’ve been practicing the way that I can tell if he’s produced something. It’s just basically if it sounds to me like he’s produced it. I know that’s annoying I will clarify first, I think it’s mostly in the vocal treatment, that’s the thing that is the most off-putting to me. But just every decision feels a little bit wrong to me. If I wanted to be cringe I could point out the Juno 6 Bass, the perfectly stacked vocals, the Juno 6 pad, the wide and tight plate reverb, the Juno 6 sword, the goo goo gaga, the man hates treble unless it’s just random crap swishing around, everything gets low passed with thin 707 drum samples, thin 808 drum samples, then 909 drum samples, thin Lin drum samples, thin thin drum samples, all the resonances, the harmonics get removed, everything is sort of squished down to the fundamental reduction in timbre, everything exists in a small frequency range, everything becomes a tone, it’s an extremely consonant sound with no interplay between the different elements. Now that is a skill to do, and it produces a cleaner sound, and the only problem is I viscerally hate that sound. ...

2. his relationship to music growing up (some gems in here)

... I got into drumming at the age of like six, so maybe as a drummer the drum samples are [offensive?]. I was anachronistically raised on cassettes so, you know maybe I like the sound of audio and maybe that explains why I tend to like dissonance and inharmonic sound, interplay of elements, distortion noise, the clashing of elements. I imagine I’ve listened to more SoundCloud in my life than radio so uh, maybe I only want to hear amateur internet losers. I’ve liked over 10 000 songs on my Spotify account and I can identify lot of them off the bat, but I’m not that great at identifying music generally, so it’s just

a compartmentalized media consumption habit. I first did music production in primary school. I made mashups and dubstep. As a teenager I extensively used Juno 6 emulators to make embarrassing synthwave. [Admittedly that's probably not the subconscious association that he's going for.] Lyrics aren't of much interest to me, it's just a sound in the mix, and so maybe I don't like the focus on vocals. I've previously described [my reaction] as a fight or flight sort of thing, so maybe it's literally is tapping into some memory of trauma. I have an unusually small ear canal so maybe I'm just hearing the high end more acutely. I was hired as a teenager to play drums on a lot of Jack Antonoff songs for a university glee club so maybe I don't like it because of the embarrassment of having ever been associated with a glee club. I liked the album 1989 but to make the song Out of the Woods palatable to me I had to mash it up with M83Midnight City and that got a DMCA takedown so maybe it's a personal vendetta. And of course I have known and know people who like this music so maybe I secretly hate my friends and family but I don't know if any of that tells us anything ...

3. An elegant and simple defense of taste

...I did use the phrase visceral hatred. But I did so only to describe my visceral hatred. I can't stand the sound of it. I hate it. And the way in which I hate it is viscerally. But I didn't go beyond that because it doesn't go beyond that. Nothing against the guy, I like Taylor Swift's stuff. It's just subjective taste...

4. His tongue in cheek and pretty based response to some of Twitter's negative response to his video

.... [it's just subjective taste]. There's just nothing to engage with, so I've had to fabricate a reason for controversy. And what I've come up with is this: **Reads off of phone** "The real reason I've been saying for years that I have a visceral reaction to Jack Antonoff's production style, the REAL reason is this: because during the rollout cycle of the album *Midnights*, I wanted to ascribe Taylor Swift's authorship to a man and diminish women's voices." So actually, I'm a sexist. So now that's out there, published as facts, in the news. With all the other facts...

5. Basically a refutation to the entire premise of my interview

.."What makes you so special and unique that you can tell the production of a song?" Nothing. ...

6. Which was refuted even harder when I found a screenshot he tweeted of his tweet drafts in June 2021, one of them reading: "people often say to me:"caleb, what's wrong with jack antonoff;'. don't ask questions"



caleb gamman
@calebgamman

...

collected unpublished jack antonoff hate 2016-2020

people often say to me: "caleb, what's wrong with jack antonoff?" don't ask questions

i don't actually have any ill will towards jack antonoff, all i'm saying is he wouldn't be able to make that music if he was dead

at the crj concert fantasising about glassing jack antonoff in the face

a lot of people have been around jack antonoff, and every single one of them has made the active decision to not kick him to the ground, spit on him and...

whenever an artist makes a genuinely good album jack antonoff pops his head around the corner "now let's do a silly one"

i will enjoy literally any piece of music except for anything

It is a perfect video essay. I pestered him anyway:

Michaela Drouillard: I know you cover this in the video. I'm sorry. What makes a Jack Antonoff song a Jack Antonoff song?

Caleb Gamman: I suppose I do have much more detail that I can go into, but it is always, for me, just a knee jerk. It's not any principled thing. I do have some loose experience with like, audio production stuff. So if I really dig into it, I can maybe point some things out. In the treble mix of the song, like in the very upper end, like sort of above where the vocals are. He often really crushes that down. Like it's lower in volume than anything else, which is sort of a weird effect. Like, when you hear someone speaking, you hear a lot of that, like, noise in their voice. When you yourself are speaking, you hear less of it, right? You hear more of your own voice bouncing around in your head. And so it sort of creates this effect of – **like he mixes his vocals the way it would sound to the person who's singing them. Which is sort of strange.** He often has no vocals on that upper end, which is very unusual, and then often there are random little bits of noise happening up there which is the sort of thing her likes to do.

So I don't know, like I personally am quite a fan of messier music, you know, more experimental stuff, stuff that has a lot of clashing, sounds, a lot of different interesting pieces coming together. Jack Antonoff does a very good job of making stuff very clean. So you know, all of the notes that are happening simultaneously are perfectly stacked. They're like mathematically stacked .. Like, you have the harmonic spectrum where everything is like a multiple on top of each other. So that's like your notes on a keyboard. And if you played a keyboard note that was a little bit off where it's supposed to be, you're gonna create this interesting effect where these two notes sort of clash against each other and they wiggle around in terms of their tone because you've got these two harmonics happening that don't match. And Jack Antonoff avoids that quite strictly.

His drums are low pass, like the vocals, so you don't get like the tinny high end in the drums or whatever because that tinny high end in the drums is like a bunch of random harmonics all over the place. So if you cut that out, you just get the tone of the drum. And then he achieves those effects using the same sounds like literally the same synthesizer on the same drum machine and stuff.

....

If you used a cscope spectrogram sort of thing to look at a Jack Antonoff song, you're gonna see a lot of big lines. And then if you're listening to a song by someone else, you're gonna see like a bunch of sort of like static wiggling. It's gonna be really jagged because you've got a bunch of harmonics that don't exactly mix with each other, so they're stacking in all sorts of weird different ways.

He was really nice and fun to talk to. I bolded the part about vocals sounding the way they would to the person who's singing them while editing this paragraph because I thought it was

kind of funny, and kind of cool. It’s coming from inside the house!!!!

2.1 So

TLDR of the Gamman interview is: Taste is subjective. People who can play instruments tend to be better at identifying and describing sounds. Cultural criticism is a hall of smoke and mirrors. Antonoff is everywhere.

Can we translate any of this into something observable? Can we quantify the Jack Antonoff-ification of pop music?

3 Data

I pulled data on artists who’ve collaborated with Antonoff using the Spotify API, which I accessed using the `spotifyr` package in R (Thompson et al. 2022). I used the `get_artist_audio_features` function to acquire data on Lorde, Taylor Swift, St. Vincent, Lana Del Rey, The Chicks, Florence and the Machine, and Bleachers discographies. I threw in HAIM, Marina and the Diamonds, Maggie Rogers, Sharon Von Etten, and Mitski in too, since they’ve never collaborated with Jack Antonoff but are more or less part of the same pop sphere.

3.0.1 Artist Selection

Of the artists who did collaborate with Antonoff, I chose these artists because they’ve produced multiple albums, and at least one album with Antonoff, as opposed to just individual songs. I chose the Chicks because I figured that collaborating with Antonoff after an extended gap in a longer spanning career than the other artists might offer some interesting results¹. The data points at beginning of their career reflect music that existed in a different time from the other artists in this dataset².

3.0.2 Data Cleaning

I joined these data sets using the `rbind` function in base R, and created an additional variable called “jack”, which contains a 1 if Jack Antonoff produced or co-produced the song, and a 0 if he had nothing to do with it. I got the information on which songs Antonoff did or did not produce from the “Jack Antonoff Production Discography” Wikipedia page (Wikipedia 2023).

¹It didn’t. But I kept them in there anyway because why not.

²Again, nothing worth noting came up.

3.0.3 Audio Features Dataset

The `get_artist_audio_features` provides a profile of each song, according to variables developed by Spotify, which makes for simple comparisons across artists. The final audio features in this study data set contains the variables: `artist_name`, `track_name`, `energy`, `danceability`, `key`, `loudness`, `mode`, `speechiness`, `acousticness`, `instrumentalness`, `liveness`, `valence`, `tempo`, `jack`.

From the `get_artist_audio_features` documentation (“Get Track’s Audio Features” (n.d.)):

- **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.
- **instrumentalness:** Predicts whether a track contains no vocals. “Ooh” and “aah” sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly “vocal”. The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.
- **key:** The key the track is in. Integers map to pitches using standard [Pitch Class notation](#). E.g. 0 = C, 1 = C / D , 2 = D, and so on. If no key was detected, the value is -1.
- **liveness:** Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.
- **loudness:** The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.

- **mode:** Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.
- **tempo:** The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.
- **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).

It’s hard to gauge exactly what the variables “danceability” or “valence” mean, even using the documentation. For instance, “valence” is described as “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)”. “Danceability” is described as “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”. Neither of these definitions fully cover what metrics and thresholds go into something like “danceability”. However, whatever these features capture is easier to compare across artists in visualizations than more granular details about bars, beats, duration, and tatums³.

This is pretty vague. Especially valence – is the positivity related to lyrics, or something else? I compared the valence scores of classical music tracks to see if this score is based on lyrics, or the just the instruments. In Table 1, Gustav Holst’s Jupiter has a valence score 326.78% higher than Mars. This tracks. It suggests the valence is based on audio features other than lyrical content. It also suggests that it is, in fact, capturing a difference between the two songs’ profiles.

Table 1: Valence and Danceability Scores for Holst’s Mars and Jupiter

Artist Name	Valence	Danceability	Track Name
Gustav Holst	0.0527	0.204	The Planets, Op. 32, H. 125: I. Mars, the Bringer of War
Gustav Holst	0.2250	0.298	The Planets, Op. 32, H. 125: IV. Jupiter, the Bringer of Jollity

I looked at the valence measures on Nikolai Rimsky-Korsakov’s Flight of the Bumblebee, a song that’s much faster and frantic than Mars and Jupiter. What I found complicated my

³-- a word I just learned, which refers to “the lowest regular pulse train that a listener intuitively infers from the timing of perceived musical events”.

understanding of valence – different performances of the same song had different valence scores (Table 2). Some were arias, played by different instruments. Arias tended to have higher valence scores. The lowest recorded valence across all performances is 0.2410, and the highest is .9230. What gives?

Table 2: Valence and Danceability Scores for Every Performance of Flight of the Bumblee on Spotify

Artist Name	Valence	Danceability	Album Release	Track Name
			Year	
Nikolai Rimsky-Korsakov	0.360	0.448	2023	Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.657	0.374	2021	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Double Reed Ensemble)
Nikolai Rimsky-Korsakov	0.742	0.476	2018	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Flute & Orchestra)
Nikolai Rimsky-Korsakov	0.534	0.509	2018	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Violin & Orchestra)
Nikolai Rimsky-Korsakov	0.742	0.476	2018	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Flute & Orchestra)
Nikolai Rimsky-Korsakov	0.534	0.509	2018	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Violin & Orchestra)
Nikolai Rimsky-Korsakov	0.241	0.429	2017	Tale of Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.923	0.663	2012	Tale of Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.241	0.429	2011	The Tale of Tsar Saltan: The Tale of Tsar Saltan: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.245	0.429	2011	The Tale of Tsar Saltan: Flight of the Bumblebee

Artist Name	Valence	Danceability	Album Release	Track Name
			Year	
Nikolai Rimsky-Korsakov	0.407	0.413	2006	The Tale of Tsar Saltan: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.359	0.226	2004	Tale of Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.510	0.425	1999	Rimsky-Korsakov: The Tale of Tsar Saltan, Act III: The Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.466	0.457	1998	The Tale of Tsar Saltan: The Flight of the Bumblebee (Arr. Rachmaninoff)
Nikolai Rimsky-Korsakov	0.444	0.440	1997	Tale of Tsar Saltan, Op. 57: Flight of the Bumblebee: Flight of the Bumble-Bee
Nikolai Rimsky-Korsakov	0.481	0.458	1995	Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.481	0.458	1995	Tsar Saltan, Op. 57: Flight of the Bumblebee

I asked Glen McDonald, a “data alchemist” (see: Principal Engineer) at Spotify who worked at The Echo Nest, a music intelligence start-up that was acquired by Spotify in 2014.

Glen told me that **valence** and **danceability** were created by giving thousands of examples to college interns and asking them to tag whether a song was positive or gloomy, or danceable or un-danceable. Variables like energy or instrumentality, on the other hand, were determined primarily through machine listening techniques. The human subjectivity part only came in when it came to tweaking the training data⁴:

“You could imagine writing a formula for energy that combines like loudness and tempo and degree of harmonic variation or something. So it was ML, but that one’s more like a human helping a machine figure out a formula. Whereas valence is teaching the machine to try to reproduce a purely human thing. Same with danceability. I mean, danceability is whether a human can dance to it.

⁴My favourite example Glen gave was that the ML kept giving bluegrass songs insanely high **speechiness** ratings. Because there were no banjos in the training data, they were registered as human speech. (This is hilarious.) They had to go back and add more songs with banjos to the training data.

So like, the machine’s not gonna dance, so the machine can’t have any opinion on that. The computer could have an opinion on energy. And the computer can definitely have an opinion on loudness. So there’s a spectrum from, you know, loudness as purely analytical, and then energy is a little like loudness, with a little more subjectivity. And then danceability and valence are purely subjective.”

This put the big range of valence scores for Flight of The Bumblebee into perspective – what that value is capturing is abstracted away from the actual notes of the song. The valence feature propagates a vibe, so of course this uncertainty makes room for noise and variation.

Glen confirmed that this process didn’t know anything about lyrics, which makes valence a particularly difficult variable to build. It could, in theory, pick up on aspects of vocal performance, but it doesn’t know anything about languages or words or the meanings of songs. Take an upbeat, happy-sounding Elliot Smith song with devastating lyrics: the machine might register it as happy, where human listeners understand that it’s sad. Add to that that two humans might even disagree on the song’s valence.

“Plus, we have the confounding factor of like, a song that seems happy today could be sad tomorrow because the singer was killed in a plane crash. The song didn’t change, but our world changed and our reactions changed.”

All this said, Glen said that the way he’s made use of these variables is by combining energy and valence to get a quadrant, which usually works “fairly well”. High energy and high valence could be generally happy, cheerful, upbeat. Low energy, low valence could be sad or downbeat. High energy, low valence is sort of angry. Low energy, high valence is serene or calming. I thanked Glen for this time, ended the call, and booted up the statistical computing programming language, R (R Core Team 2022).

In Figure 1, it looks like songs produced by Jack Antonoff are a little more clustered towards the lower end of valence, and across the spectrum of energy, although the variance is pretty big.

In Figure 2, it looks like Jack-produced songs err towards higher danceability, across the spectrum of energy.

These quadrants are a start, but they don’t suggest any defining Jack-iness characteristic underlying each of those blue points, which are scattered pretty broadly.

By comparing cumulative distribution functions of each variable based on whether or not a song was produced by Antonoff, we can see that some of these variables do capture some sort of underlying Jack-iness (Figure 3). For instance, while liveness, instrumentalness, and speechiness are relatively similar, we can see clear differences in danceability, energy, loudness, acousticness, and valence.

Another way to gauge the Antonoff-ification of pop music could be to look at how artists’ sounds have changed over time, and whether or not there is any correlation with Antonoff. Here is where the necessary “correlation is not causation” caveat comes in: suggesting that

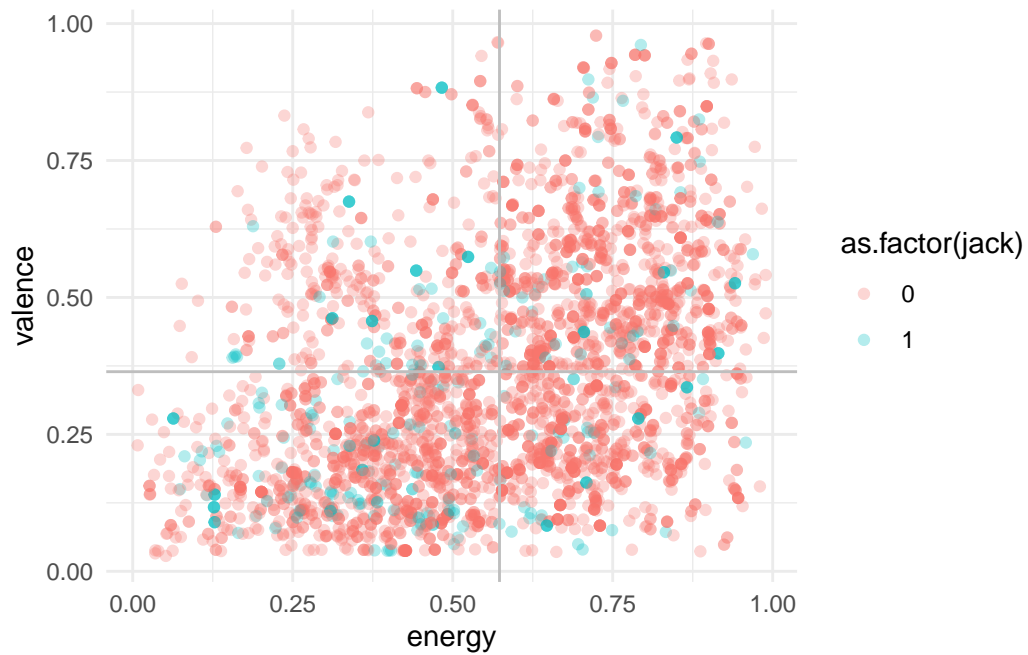


Figure 1: Valence and Energy Quadrant

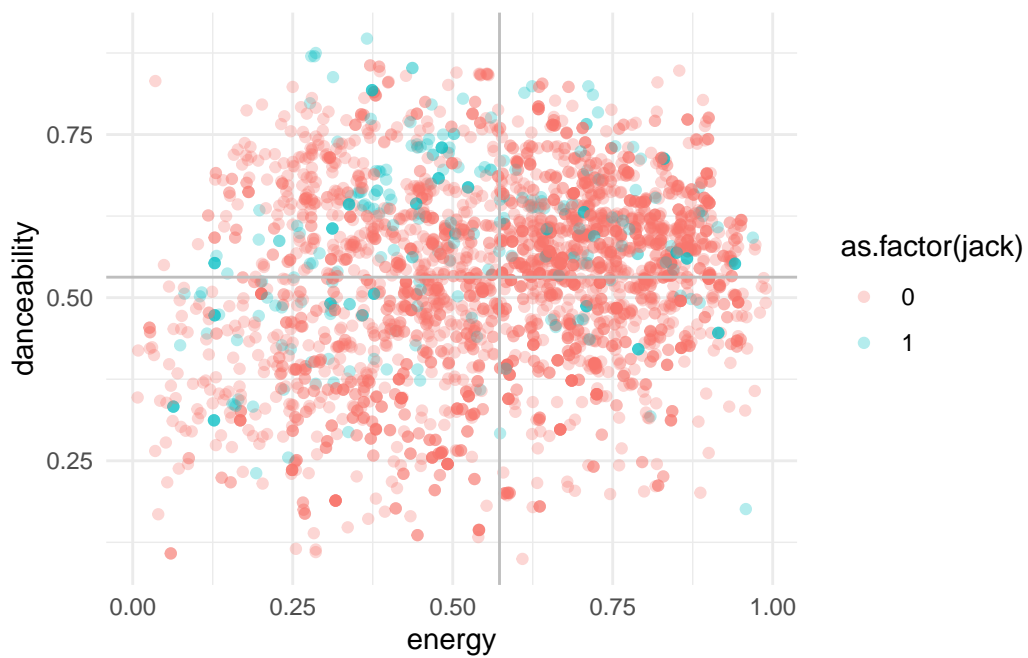


Figure 2: Danceability and Energy Quadrant

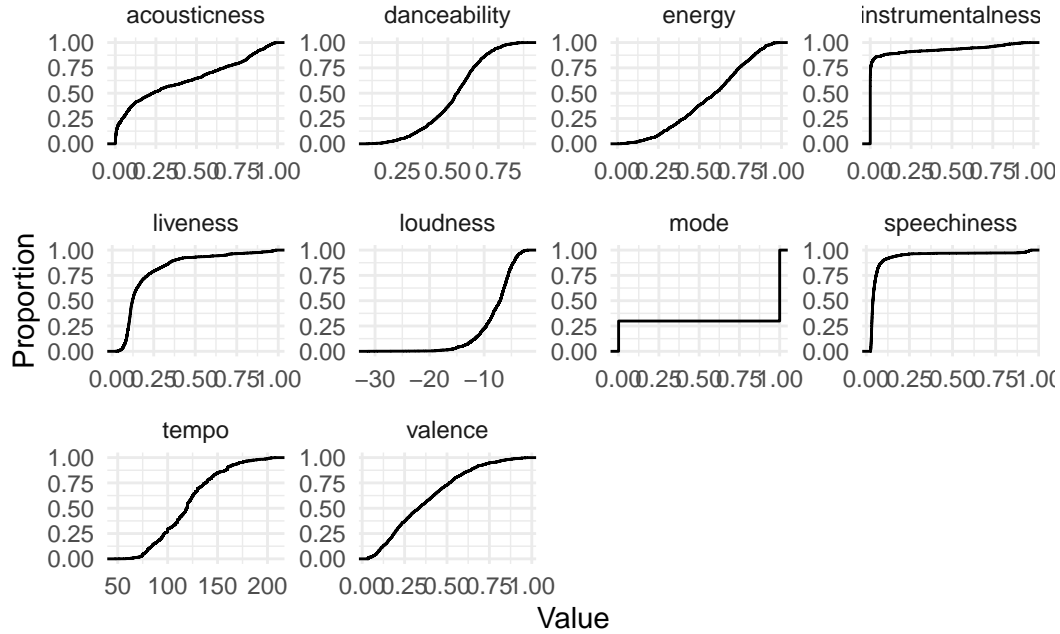


Figure 3: Cumulative Distribution Function of Each Variable,

Antonoff is the *cause* for any change in artists' sounds would be an overstep. These plots are here to get a glimpse at pop's evolution over time, and where Antonoff fits in that picture. Each point represents a song.

In Figure 4, it looks like there's a downward trend in energy scores in artists who work with Antonoff, but that these downward trends are also somewhat in line with the artists' energy trajectories anyway. Specifically, Lana Del Rey and Taylor Swift. They have enough albums recorded prior to collaborations with Antonoff to be able to observe this. Sharon Von Etten, Maggie Rogers, and Mitski, who have never worked with Antonoff, all have upward trends.

There aren't any significant correlations in danceability (Figure 5) .

In Figure 6, it looks like Lorde has the only significant change in valence scores, which tracks with my understanding her albums. Melodrama had Green Lights, and Solar Power had a music video of Lorde frolicking on a beach in a yellow bikini with Phoebe Bridgers and Clairo.

Again in Figure 7, Lorde's acoustic Solar Power album jumps out at us. Taylor Swift, Lorde and Haim also have significant-seeming upward trends in acousticness, while Mitski, Sharon Von Etten and Maggie Rogers see decreasing acousticness.

In Figure 8, St. Vincent is the only plot with any significant trend in instrumentalness.

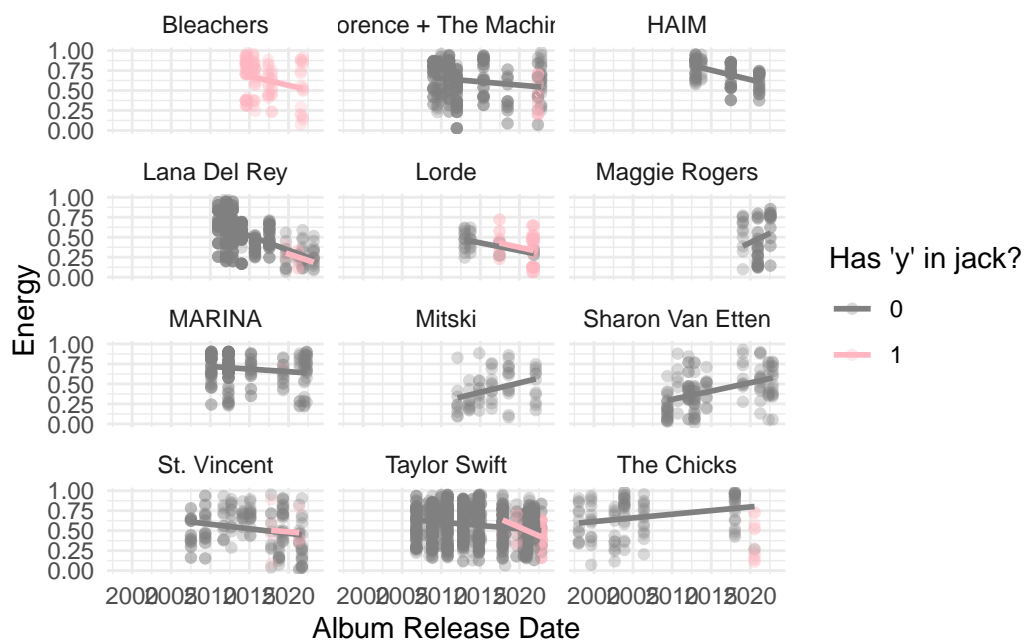


Figure 4: Energy Variable

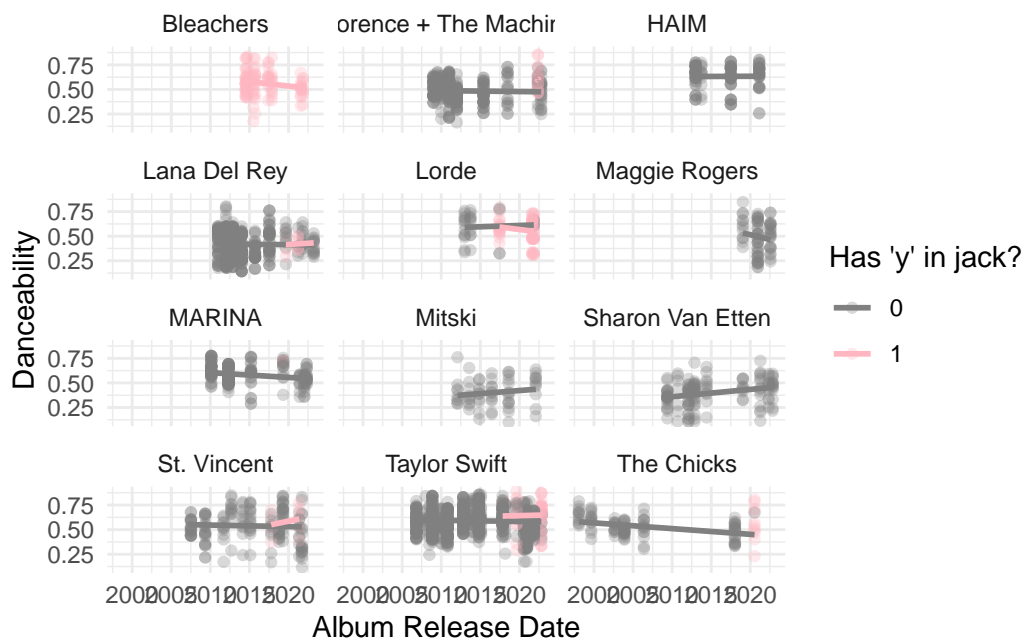


Figure 5: Danceability Variable

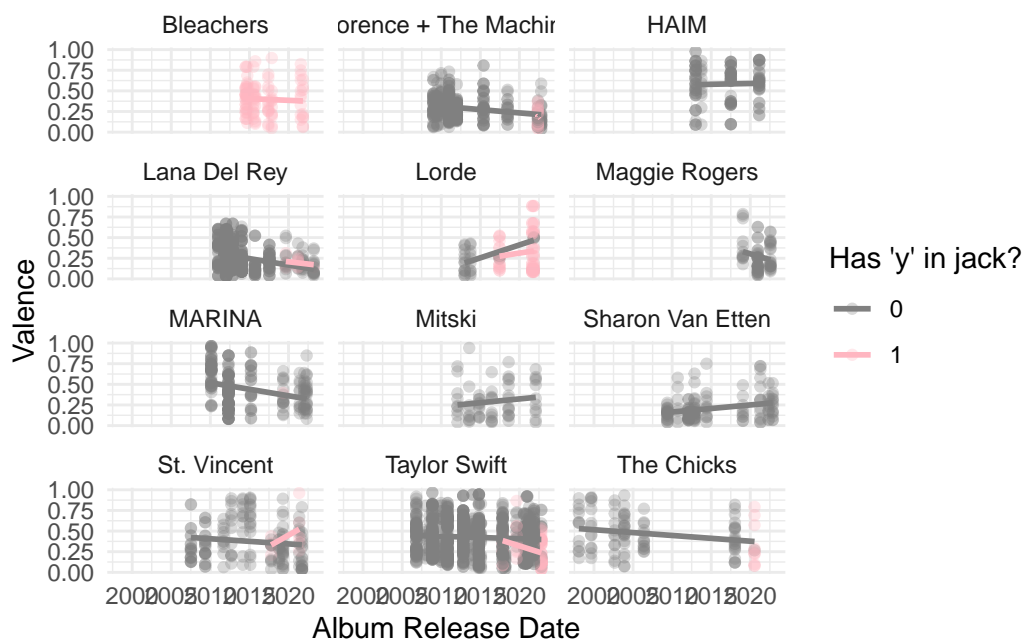


Figure 6: Valence Variable

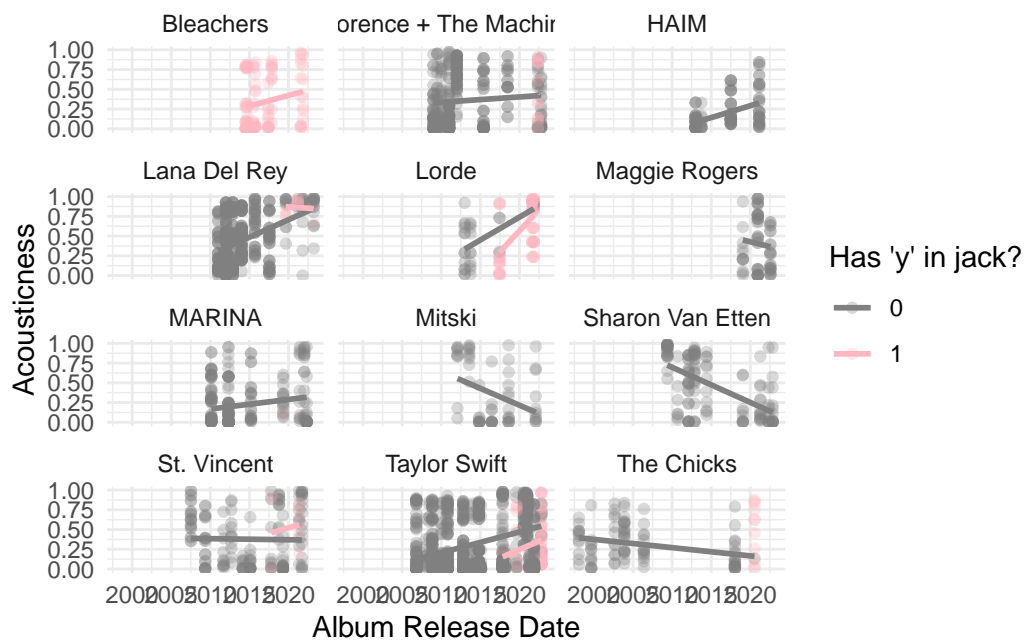


Figure 7: Acousticness Variable

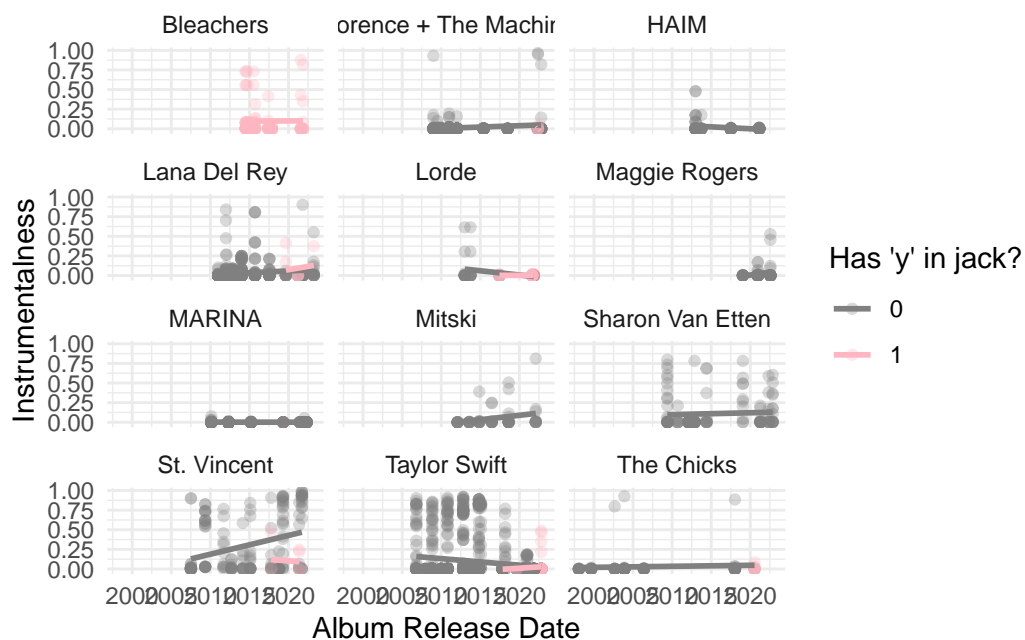


Figure 8: Instrumentalness Variable

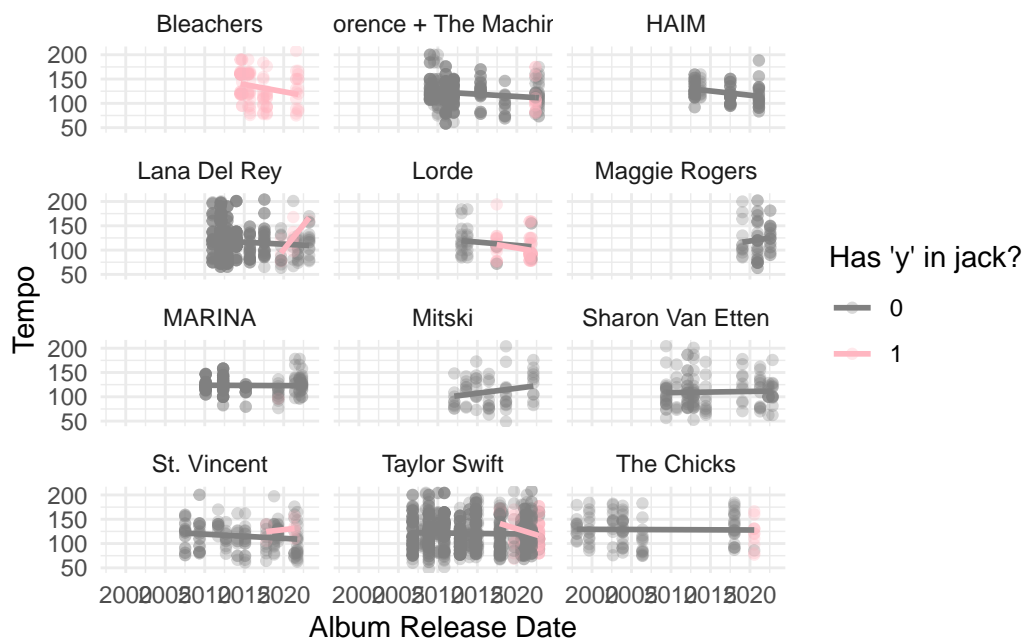


Figure 9: Tempo Variable

Finally, in Figure 9, there aren't many significant trends in tempo, except for Antonoff's collaborations with Lana Del Rey and Taylor Swift.

Now that we understand that there do seem to be some subtle differences in how variables that are and aren't produced by Jack behave, as seen in our cumulative distribution functions, our quadrants, and some minor trends in artists' sounds over time (specifically, **energy**), we can use a logistic regression to see if a model can a) predict whether or not a song was produced by Antonoff based on these variables, and b) tell us which variables were more significant in determining the output.

4 Model

Since we're trying to figure out whether these variables can tell us anything about what Jack Antonoff's sound is all about, a good way to figure that out is by seeing if any of these variables can help us predict what songs are produced by Jack. Logistic regressions like questions with binary outcomes.

I made a binary response variable, called **jack**, that has a value of 1 if the song was written by Jack, and a 0 otherwise. There are 10 predictor variables, all of which are continuous: **danceability**, **energy**, **loudness**, **mode**, **speechiness**, **acousticness**, **instrumentalness**, **liveness**, **valence**, and **tempo**.

The formula for this logistic regression can be written as:

$$\Pr(y_i = 1) = \text{logit}^{-1}(\beta_0 + \beta_1 x_i)$$

where $\Pr(y_i = 1)$ is the probability of a song being written by Jack, β_0 is the intercept, and $\beta_1, \dots, \beta_{10}$ are the coefficients corresponding to the predictor variables.

Figure 10 gives us an idea of the mean value of each of these features for Jack and non-Jack song. A logistic regression will give us a better idea of how to weigh each of these features in order to predict the outcome, which gives us information about the "research" "question": are these characteristics related to Jack's music? (And then: Which ones? And then, the most elusive question: How?)

Bracketing the major blindspot that these variables provide a single number to represent the dynamics of an entire song, the model has a few other limitations. First of all, this model assumes that the relationship between the independent variables and the dependent variable is linear. There could be non-linear relationships between the variables that this model isn't capturing. Second, it assumes that the relationship between the independent variables and the dependent variable is the same for all observations, which might not be true in practice – there could be interactions between the variables, or in certain subgroups of data where the relationship is different. Albums where Jack may have had more of a back seat producing

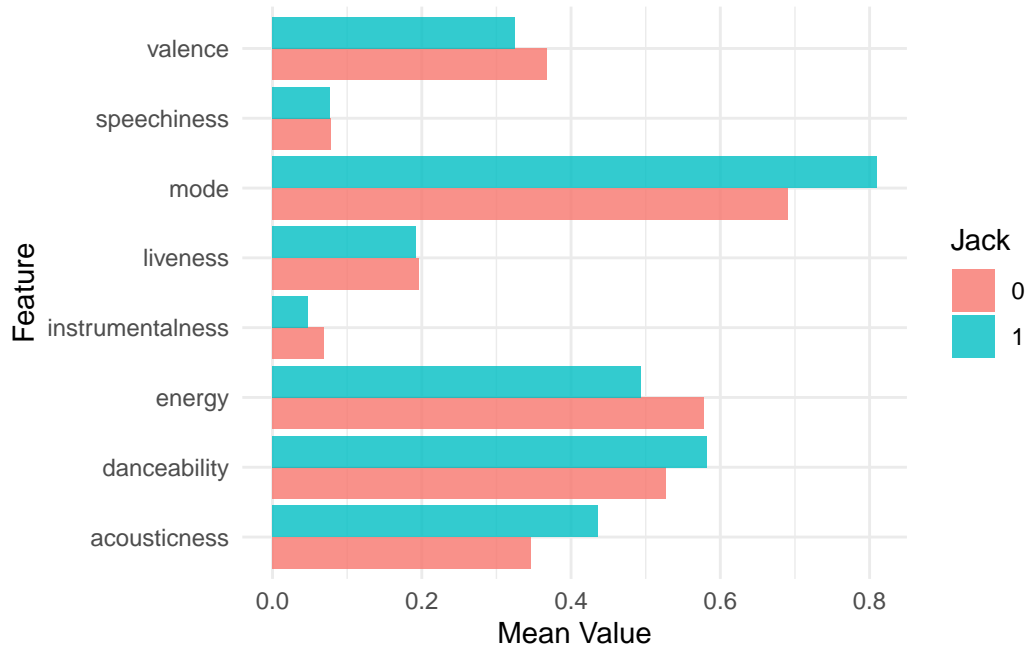


Figure 10: Comparing the Mean Values of Each Variable

role are one example. (The fact that a song’s development, from inception to recording, is a whole creative process, each having its own origin story with varying degrees of Jack-ness, is another.)

One issue is that in our dataset, there are a lot more songs that aren’t produced by Jack than songs that are:

Is it Jack?	Number of Songs
0	3110
1	294

So, we have to balance the classes.

After some minor pre-processing using the **tidyverse** R package (Wickham et al. 2019). I followed Julia Silge’s tutorial (Silge 2023) to build a model that deals with the imbalanced classes.

I used the **tidymodels** to fit the model (Kuhn and Wickham 2020). To deal with the imbalanced classes, I downsampled the majority class using the **step_downsample()** function from the **themis** package. Then I used 5-fold cross-validation with stratification by the **jack** variable to estimate the performance of the model.

The `workflow_set()` function lets us hold different sets of feature engineering together and use tuning to evaluate them all at once. I need to evaluate two elements: a basic model, and a model that was downsampled using the `themis` package.

I used `rank_results()` from the `yardstick` package to rank the models based on accuracy, log loss, sensitivity, and specificity, and went with the downsampled model.

5 Results

The model's accuracy score is 0.743, and the ROC AUC score is 0.841. The accuracy score could be misleading, since the model had imbalanced classes, and downsampling may have affected the model's ability to correctly identify the minority class. The ROC AUC score is more reliable, since it's less sensitive to class imbalance.

Table 4: Confusion Matrix for Logistic Regression Results

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	581	25
Predicted Negative (0)	194	51

Table 5 shows the different variables' importance on determining whether a song was had any Antonoff involvement.

Table 5: Importance Scores of Each Variable in Determining Jack-ness

Variable	Importance	Sign
danceability	4.3200694	POS
instrumentalness	2.5899482	NEG
valence	2.0246212	NEG
speechiness	1.7422846	NEG
energy	1.3839382	POS
liveness	1.0386332	POS
mode	0.5269857	POS
acousticness	0.3696266	NEG
loudness	0.3531411	NEG
tempo	0.0072614	POS

Danceability, with a positive importance score of 4.32, had the highest level of importance in determining whether or not a song was produced or co-produced by Jack Antonoff. Instrumentalness comes in second at 2.59, with a negative score (meaning that negative instrumentalness is correlated), and valence comes in third with a negative score at 2.02.

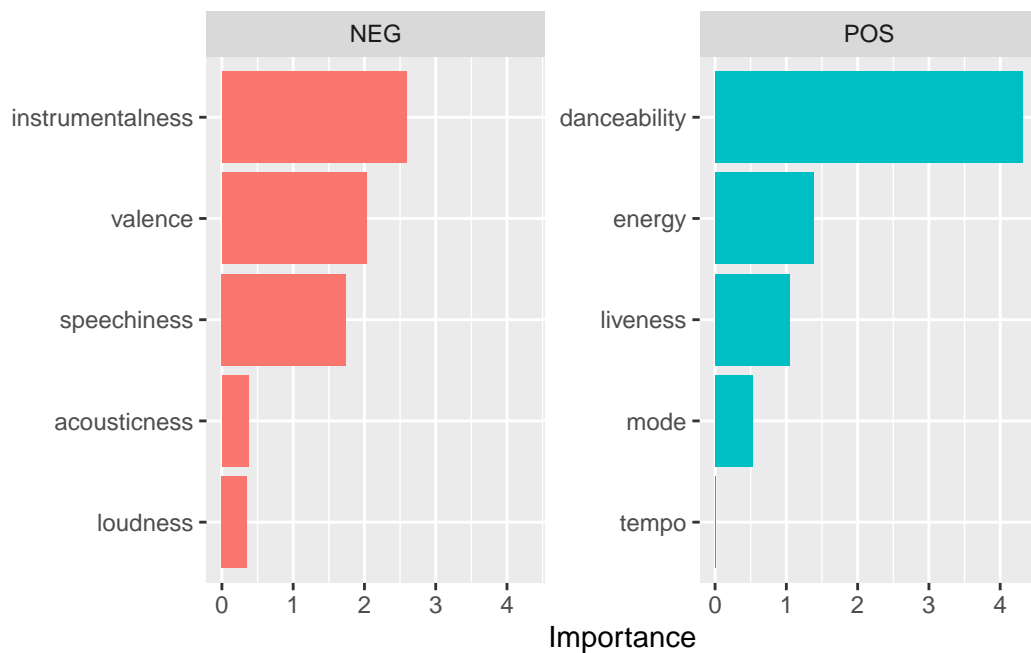


Figure 11: Importance Scores of Each Variable in Determining Jack-ness

Table 6: Top 20 Most Danceable Antonoff-Produced Songs

Artist Name	Danceability	Track
Taylor Swift	0.897	I Think He Knows
Taylor Swift	0.875	Vigilante Shit
Florence + The Machine	0.852	Heaven Is Here
Bleachers	0.838	I Wanna Get Better - Demo Version
Taylor Swift	0.824	Cornelia Street
Bleachers	0.824	Hate That You Know Me - MTV Unplugged
Bleachers	0.818	Wake Me
Bleachers	0.814	Shadow
Taylor Swift	0.811	Paper Rings
The Chicks	0.805	Julianna Calm Down
Lorde	0.796	Sober
The Chicks	0.784	Texas Man
Bleachers	0.776	Hate That You Know Me
St. Vincent	0.774	Daddy's Home
Bleachers	0.769	All My Heroes
Taylor Swift	0.766	Look What You Made Me Do
Taylor Swift	0.751	Question...?

Artist Name	Danceability	Track
Taylor Swift	0.743	Lavender Haze
Taylor Swift	0.739	False God
Bleachers	0.732	Goodbye

Table 6 shows us the 20 songs in Jack Antonoff’s discography with the highest danceability scores. Bleachers takes up 35% of this list, and Taylor Swift takes up 40% of this list. St Vincent, The Chicks, Lorde, and Florence and the Machine also have songs up there.

6 Discussion

6.1 Danceability Predicts Jack-iness

Using the Spotify Audio Features, you can’t really gauge what Jack Antonoff’s music *sounds* like. You get a rough idea, but you don’t get the same granularity that you would by simply DM’ing Caleb Gamman on Twitter and asking him to describe it to you. Or by watching the video where he already did that.

I think it’s mostly in the vocal treatment, that’s the thing that is the most off-putting to me. But just every decision feels a little bit wrong to me. If I wanted to be cringe I could point out the Juno 6 Bass, the perfectly stacked vocals, the Juno 6 pad, the wide and tight plate reverb, the Juno 6 sword, the goo goo gaga, the man hates treble unless it’s just random crap swishing around, everything gets low passed with thin 707 drum samples, thin 808 drum samples, then 909 drum samples, thin Lin drum samples, thin thin drum samples, all the resonances, the harmonics get removed, everything is sort of squished down to the fundamental reduction in timbre, everything exists in a small frequency range, everything becomes a tone, it’s an extremely consonant sound with no interplay between the different elements. Now that is a skill to do, and it produces a cleaner sound, and the only problem is I viscerally hate that sound.

The Spotify Features also wouldn’t help you understand in the same way you end up understanding it if you sat on my friend Matthew Jordan’s couch and watched him beat box a few different Jack Antonoff and Pharrell impressions in order to show the contrast.

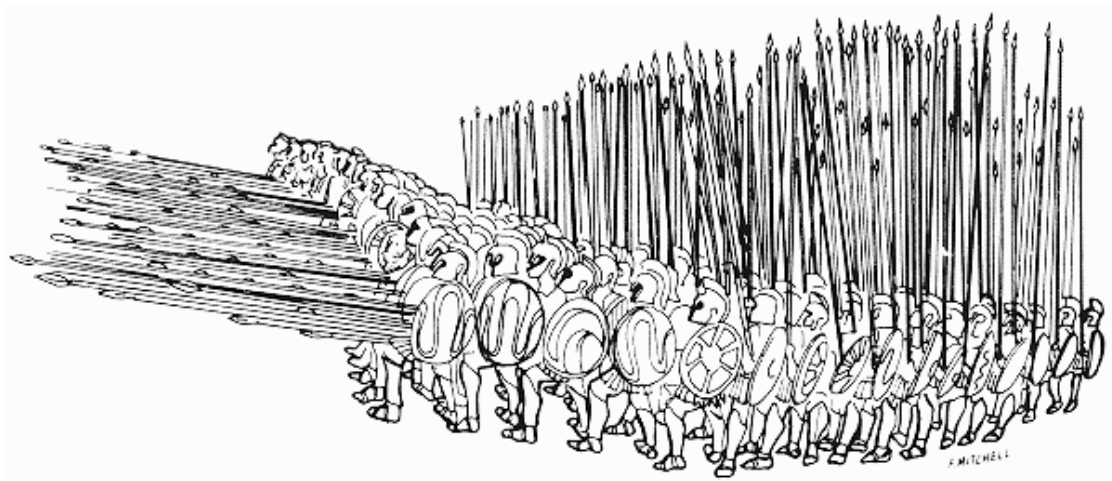
According to this model, though, danceability is the most important feature in predicting whether or not a song was produced by Jack Antonoff. And the accuracy metrics aren’t awful – the accuracy score is 0.743, and the roc_auc (receiver operating characteristic area under the curve) (more fun to say out loud as an abbreviation. Rock Auck), which is a metric better suited for imbalanced datasets, such as our own, is 0.810. This means that

although danceability can't show us the inside of a song, something about this score is telling us something about Jack's music.

This was a feature that was invented in a lab ~15 years ago. This number refers to the time a bunch of undergraduates were paid (hopefully?) to listen to a bunch of songs and tag whether they were danceable or not. It is pretty much based on intuition. It captures this je-ne-sais-quoi essence of music that connects with listeners on a visceral level. Even if it doesn't make YOUR body want to move, you could imagine it making SOMEBODY'S body want to move.

6.1.1 In which I do a close listen of our model's top 5 most danceable songs⁵:

- I Think He Knows by Taylor Swift:
 - Pretty danceable tbh.
- Vigilante Shit by Taylor Swift:
 - Not danceable. But it is trying to coerce you into some sort of body feeling.
- Heaven is Here by Florence and the Machine
 - DEFINITELY not danceable. But it is trying to get your body to do something. Kind of sounds like it's suggesting that you join a Macedonian phalanx or something. But no. Not danceable.



- I Wanna Get Better by Bleachers

⁵I encourage you to follow along and drop me an email with your own assessment at michaela.drouillard@mail.utoronto.ca.

- I don’t know if “danceability” is the word, but I’ll give it “anthem”-y. Like, I can’t picture a group of friends dancing to this in the kitchen, but I can picture someone driving, on their way to give a presentation or negotiate a promotion or something, and this song comes on the radio and they leave it on. They kind of lean into it.
- Cornelia Street by Taylor Swift
 - Yeah I’m dancing. Or at least bobbing my head. I only became aware of that fact when I made eye contact with the guy next to me at this cafe.

6.2 Danceability isn’t even necessarily about Dancing! I propose: Visceral-ness

I am thinking about what Gamman said in his video.

I did use the phrase “visceral hatred”. But i did so only to describe my visceral hatred. I can’t stand the sound of it. I hate it. And the way in which I hate it is viscerally. But I didn’t go beyond that because it doesn’t go beyond that. Nothing against the guy, I like Taylor swift’s stuff. It’s just subjective taste.

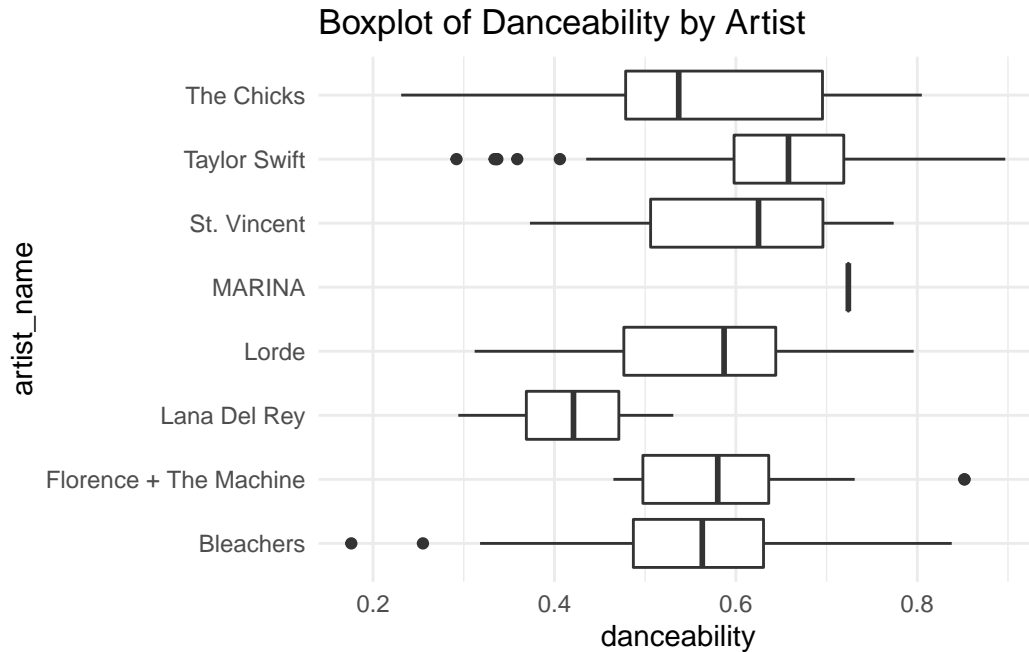
It’s just his subjective taste, but in one way Gamman was actually totally on the money about this. The danceability variable, which is the strongest predictor of Jack-ness, seems to actually be more about “visceral”-ness when you do a close-listen of the songs with the highest danceability scores. Not all of them are danceable, but the one’s that aren’t still feel like they’re trying to manipulate the way your body is holding itself in some way. The reason Jack is getting under people’s skin could just be because his music... gets underneath your skin. Right to the viscera.

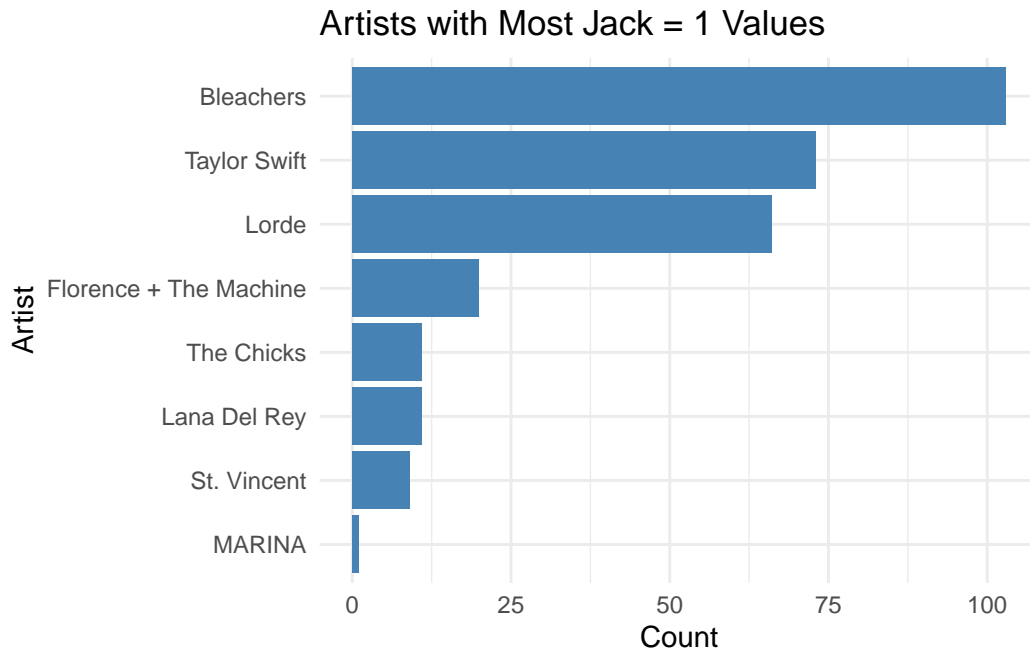
6.3 Weaknesses and next steps

One fundamental weakness is cooked into the dataset. Since we’re not in the studio, we can’t know the degree of Jack-iness in each individual song. Especially when he takes on the role of a more backseat co-producer. We can’t actually know what his involvement in each song was. There’s also the issue that a study like this implies that Jack Antonoff is the reason for the way these artists’ sounds develop, which takes that agency away from the artists. This is a critique you see on the internet, and also a critique that some people on Twitter leveraged against Gamman after his Midnights video went viral. I mentioned it to a friend when she asked me how my final paper was going, and she said something like: “Yeah but like, it’s *Taylor Swift*. As if anyone actually thinks that Jack Antonoff is the reason for her success.” I agree.

6.3.0.1 Could our model just be predicting Jack via Taylor Swift?

So another issue is that there is a lot of Taylor Swift in our training data. As mentioned earlier, her songs take up 40% of Table 6, which is 5% more than Antonoff's own band, Bleachers. Taylor Swift has the second highest average danceability score across all songs, and the most danceable songs are hers. She also happens to have a lot of collaborations with Antonoff under her belt – more than any other artist. I suspect that perhaps our model was predicting Jack via Taylor, which could be a beautiful story about collaboration and aesthetic affinity. Or it could suggest that this model has a stink to it. An avenue for future study.





6.3.0.2 Next Steps

It only makes sense, of course, to hire a bunch of undergrads and get them to tell us whether or not a few thousand songs are visceral, and then build our own visceral-ness metric, and see if we get the same results – that danceability is just visceral-ness in a trench coat, and that visceral-ness is the characteristic that helps us best understand what makes Jack, Jack.

And if we're doing this for danceability, there might also be something worth investigating in valence, the feature that indicates a song's "positivity", which was also built via college interns' intuitions.

References

- @calebgamman. 2022. “Im Able to Instantly Detect If Jack Antonoff Worked on a Song Due to a Visceral Hatred of His Production Style.” <https://twitter.com/calebgamman/status/15834747267979649>.
- Gamman, Caleb. 2022. *The Internet’s Top Jack Antonoff Hater DESTROYED*. https://www.youtube.com/watch?v=cG175pQYMxA&ab_channel=calebgamman.
- “Get Track’s Audio Features.” n.d. <https://developer.spotify.com/documentation/web-api/reference/get-audio-features>.
- Kuhn, Max, and Hadley Wickham. 2020. *Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles*. <https://www.tidymodels.org>.
- Marantz, Andrew. 2022. “Jack Antonoff’s Pop-Music Collaboration Machine.” *The New Yorker*, May. <https://www.newyorker.com/magazine/2022/05/23/jack-antonoff-pop-music-collaboration-lorde-taylor-swift>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Silge, Julia. 2023. “To Downsample Imbalanced Data or Not, with #TidyTuesday Bird Feeders.” Julia Silge. 2023. <https://juliasilge.com/blog/project-feederwatch/>.
- Thompson, Charlie, Daniel Antal, Josiah Parry, Donal Phipps, and Tom Wolff. 2022. *Spotifyr: R Wrapper for the ‘Spotify’ Web API*. <https://CRAN.R-project.org/package=spotifyr>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wikipedia. 2023. “Jack Antonoff Production Discography — Wikipedia, the Free Encyclopedia.” https://en.wikipedia.org/wiki/Jack_Antonoff_production_discography.